Search for a massive resonance decaying to qV/VV/VH in hadronic final states

by

Yongjie Xin

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

May, 2015

Abstract

Two searches for massive resonances decaying into qV/VV/VH final states are presented, in which V denotes W, Z bosons, and H denotes Higgs boson.

The data analyzed corresponds to an integrated luminosity of 19.7 fb^{-1} collected by the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. The cutting edge jet substructure algorithms are used to distinguish V and H boson jets from the standard model QCD quark/gluon jets background.

Exclusion limits are set at a confidence level of 95% on the production cross sections of: (i) excited quark resonances q* decaying to qW and qZ for masses less than 3.2 TeV and 2.9 TeV, respectively, (ii) a Randall–Sundrum graviton G_{RS} decaying into WW for masses below 1.2 TeV, and (iii) a heavy partner of the W boson W$'$ decaying into WZ for masses less than 1.7 TeV. In HVT model scenario B, resonance masses are excluded for W$'$ in the interval [1.0, 1.6] TeV, for Z$'$ in the intervals [1.0, 1.1] and [1.3,1.5] TeV, and for mass-degenerate W$'$ and Z$'$ in the interval [1.0, 1.7] TeV. A model with a bulk graviton G_{Bulk} that decays into WW or ZZ bosons is also studied.

ABSTRACT

Primary Reader: Petar Maksimovic

Acknowledgments

I would like to thank my Ph.D advisor, Petar Maksimovic. He taught me how to accomplish a physics analysis. He taught me how to collaborate with people in a large organization. He taught me how to stay calm when bad things happen. He shared with me his life stories to help me survive in such a different culture. He is always so supportive and always happy to help. I would like to express my sincere gratitude for him being a brilliant advisor and a fantastic mentor.

There are also many people to thank for my education in experimental particle physics, but probably none more than Andreas Hinzmann. I thank him for teaching me all the techniques and always being super patient in answering my questions. I thank him for being a wonderful leader and an excellent colleague.

I would like to thank Professor Morris Swartz for helping me with analysis questions and for being my analysis committee and thesis committee.

I would like to thank my colleagues and friends in the physics department. Thanks for Lei Feng and Can You for being wonderful officemates, discussing the analysis questions and sharing frustrations and happiness as being a physics Ph.D. Thanks

ACKNOWLEDGMENTS

for Guofan Hu for his helps in these years, and for being my student mentor for the first two years of my Ph.D life. Thanks for all the advices and suggestions. I would like to thank David Fehling, Kevin Nash, Marc Osherson, Ian Anderson, Chris Martin for helping me with specific analysis questions and for being great colleagues for all these years. I would like to thank my JHU friends, Longyu Zhao, Yuchong Shao, Kui Chen, Zhilei Xu, Bingxiao Xu, Yaofu Zhou, Jiajia Wen, Shan Wu, Danru Qu, Ting Su, Liang Dai, Hongbin Chen and etc.

I would like to thank my parents, for their love and care, for their support and understanding, for their patience and guidance. Last but not the least, I would like to thank my fiancee, Tian Mao, for her support and for sharing life and happiness with me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Standard Model

A physics model is a description of a system using physics and mathematical concepts and language. Standard Model (SM) is a physics model that summarizes what is currently known about the subatomic world. The SM is composed of two gauge symmetry theories: the Glashow-Salam-Weinberg Model (GSW) [1–3] describing electroweak interactions and the Quantum Chromodynamics (QCD) [4] describing the strong interactions. In SM, there are two types of fundamental particles, fermions and bosons. Fermions have half integer spin and are the building block of matter. Bosons have integer spin and are the intermediate particles of the four fundamental interactions of SM: strong, weak, electromagnetic, and gravitational force. In this chapter, we briefly introduce the components of the SM and the related theories to

this thesis.

1.1.1 Fundamental particles

Fermions are de ned as particles with half integer spin (intrinsic angular momentum), like leptons, quarks, proton, etc, while bosons are particles with integer spin, like the photon and the Higgs boson. The fundamental particles of SM are shown in Figure 1.1.

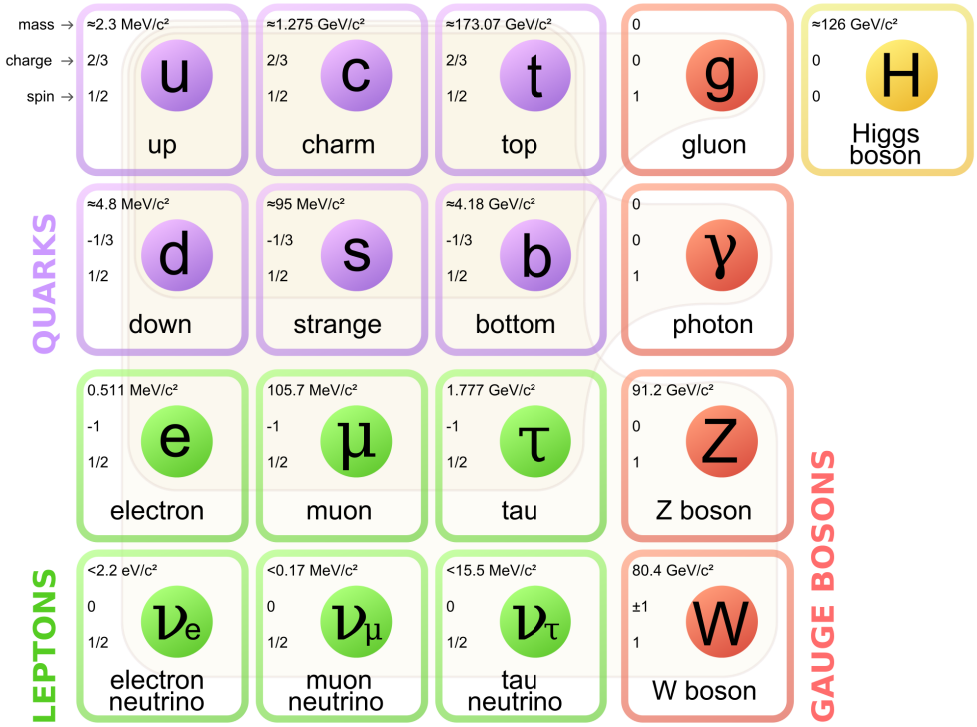

Figure 1.1: Standard model of elementary particles: the 12 fundamental fermions and 5 fundamental bosons. [5]

Quarks

In SM, there are six avors of quarks: up (u), down (d), strange (s), charm (c),

top (t), and bottom (b). Quarks carry properties like flavor, color, spin, charge, mass, etc. All quarks have spin $\frac{1}{2}$ in SM and they belong to one of the three generations: u and d quarks belong to the first generation; s and c quarks are the second generation quarks; t and b quarks make up the third generation. In Quantum Chromodynamics, color is conserved in interactions. There are 3 types of color: red, blue, green. A particle is colorless if it carries a net color charge of zero.

Leptons

Leptons are fundamental particles, which are also fermions. As shown in Figure 1.1, there are three generation of leptons: electron (e) and electron neutrino (ν_e) are the first generation; muon (μ) and muon neutrino (ν_μ) are the second generation; tau (τ) and tau neutrino (ν_τ) are the third generation. Leptons have electric charge, but do not carry color charge. Thus they are involved in the electroweak interactions, but not the strong interactions.

Bosons

Every interaction has its mediator: the photon (γ) for electromagnetic force, W and Z boson for weak force, gluon (g) for strong force and graviton (not found yet) for gravity. Higgs boson (H), although not a mediator, accounts for the mass of other fundamental particles, which will be elaborated in the following section. The W, Z, γ and g bosons having spin 1, are called vector boson. H boson has spin 0, which is called a scaler boson. There are two W bosons, distinguished by their electron charges, W^+ and W^-.

1.2 Fundamental interactions

The four fundamental interactions are summarized in Table 1.1. The "Strength" column is the relative effective strength of these forces.

Since gravity is so small compared to other three forces, it is mostly not considered in the process of particle physics. The strong force is described by the Quantum Chromodynamics. And the electromagnetic force and weak force are unified into electroweak interaction, described by the Glashow-Weinberg-Salam (GSW) model. The theory of particular interest to this thesis is the Quantum Chromodynamics (QCD), which will be introduced in the following section.

Force	Strength	Mediator
Strong	10	Gluon
Electromagnetic	10^{-2}	Photon
Weak	10^{-13}	W and Z
Gravitational	10^{-42}	Graviton

Table 1.1: Summary of the four fundamental forces in Standard Model [6].

1.2.1 Quantum Electrodynamics (QED)

QED is the theory describing the electromagnetic interaction: the interaction between electric charged particles via the exchange of photons. The "interaction vertex"

in QED, in the form of Feynman diagram, is shown in Figure 1.2. An electron comes in, and radiates a photon, then goes out. Feynman diagrams are pictorial representations of the mathematical expressions describing the interactions of elementary particles [7]. But it can be interpreted as the amplitude of a process. By connecting couple of these primitive processes together, we can get complex process like the one in Figure 1.3. In Figure 1.3, the electron and positron annihilate into a photon, which further decays to a pair of quarks, and then one of the two quarks radiates a gluon. The photon in Figure 1.3 is a virtual particle. Virtual particles are not observable, which are called o -shell particles. While real particles, in Figure 1.3, are the incoming electron and positron, and outgoing quarks and gluons, which can be observed and are on-shell .

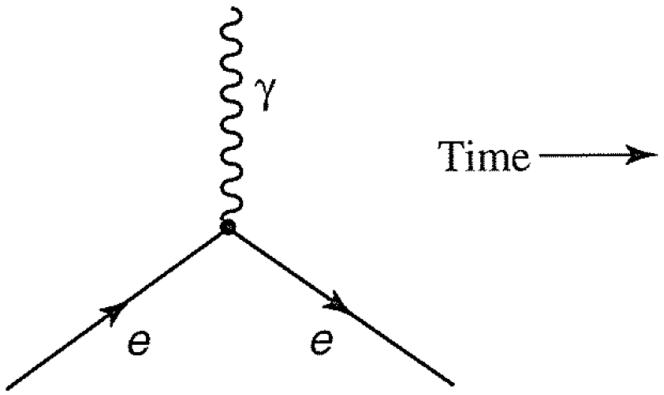

Figure 1.2: The primitive QED process in SM.

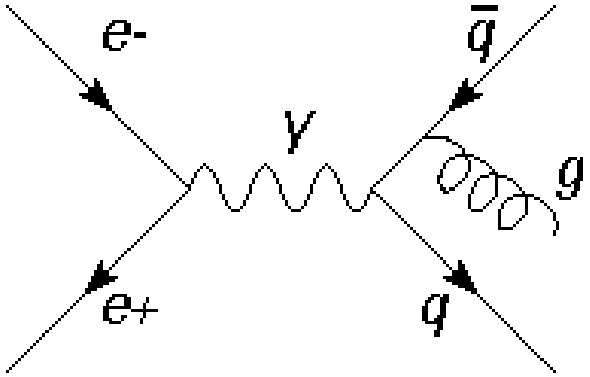

Figure 1.3: Feynman diagram for electron annihilation.

1.2.2 Quantum Chromodynamics (QCD) and jets

QCD is a theory describing the strong force and the involved fundamental particles. As we see from Table 1.1, strong force is the strongest force of the four. However, unlike the long range electromagnetic force, the strong force could only affect $\approx 10^{-15}$ m (1 fm), which is about the radius of a nucleus.

Color is one of the unique properties of QCD. The color of QCD is an analogue of the electric charge of QED. There are three types of colors charges: red, blue, green. Each quark carry one kind of color. So, for example, there are three types of top quarks in QCD: the blue top quark, the green top quark, and the red top quark. This is the same for all the other five flavors of quarks. Anti-quark carries one kind of anti-color. Gluons is the boson intermediating the strong force, just like the photon is the mediator of the electromagnetic force. Gluons carry a color and an anti-color. When two quarks interact with each other, they interact through a gluon by exchanging colors. For example, as shown in Figure 1.4, one red quark comes in,

radiates a gluon with color red and anti-blue, and a blue quark goes out. In terms of the color SU(3) symmetry, there are 8 gluons in QCD, as shown in Equation 1.1.

$$
\begin{aligned}
g_1 &= (rb + br) \quad \bar{2} & g_2 &= \quad i(rg \quad gr) \quad \bar{2} \\
g_3 &= \quad i(rb \quad br) \quad \bar{2} & g_4 &= (bg + gb) \quad \bar{2} \\
g_5 &= (rr \quad bb) \quad \bar{2} & g_6 &= \quad i(bg \quad gb) \quad \bar{2} \\
g_7 &= (rg + gr) \quad \bar{2} & g_8 &= (rr + bb \quad 2gg) \quad \bar{6}
\end{aligned}
\tag{1.1}
$$

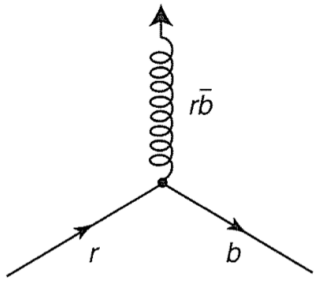

Figure 1.4: The illustration of quark-gluon interaction [6].

Unlike charged particles, colored particle could not be free. What we see in experiments and daily life are color-singlets. So we can never observe a single quark or a single gluon in an experiment. What we observe is the hadronization products of

quarks and gluons, which are called "jets" and introduced in the following text. The existence of quarks and gluons is proved by indirect experiments via jets. Here we introduce the essential components of QCD.

The coupling constant

The coupling constant is a scaler quantity describing the strength of an interaction. Here in QCD, the coupling constant α_s is :

$$\alpha_s(|q^2|) = \frac{2\pi}{(11n - 2f)ln(|q^2|/\Lambda^2)} \quad (|q^2| \gg \Lambda^2) \tag{1.2}$$

where $|q|^2$ is the squared energy-momentum 4-vector of the mediator gluon, n is the number of colors (3, in SM), f is the number of flavors (3, in SM), and Λ is a parameter determined from experimental data, which is in the range of 100∼500 MeV.

Note that α_s is not a constant. It is a function of $|q|^2$, the interaction energy. Thus it is also named as the "running coupling constant". The experimental measurements are shown in Figure 1.5.

Asymptotic freedom

As we see from Equation 1.2, when $|q|^2$ increases, α_s decreases. At large $|q|^2$, corresponding to short distances ($\leq 1\ fm$), the strong force is so weak that quarks inside of proton travel freely. At very high energies, it is also possible to form quark-gluon plasma, since their interactions are so weak. This is called the asymptotic freedom.

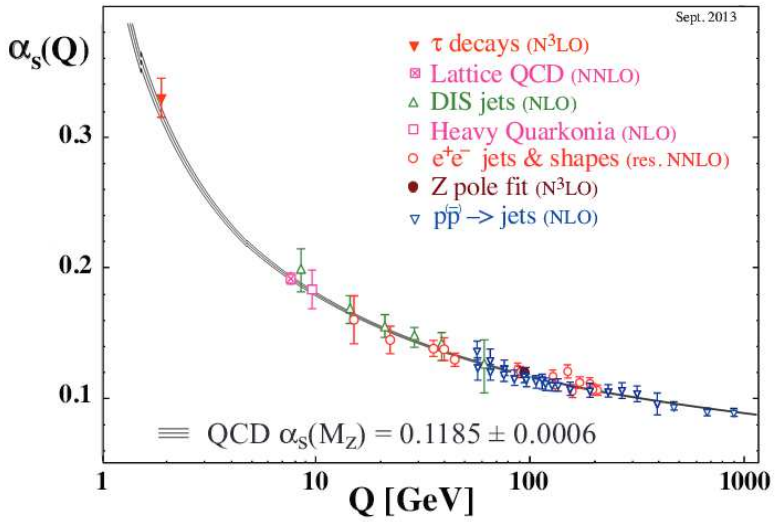

Figure 1.5: The running coupling constant of QCD. NLO is short for the next leading order. NNLO is short for next-next-leading order.

QCD confinement

QCD confinement means that the force between quarks largely increases when they are separated. So one has to exert a lot energy to try to separate a quark from other quarks. And this energy will become large enough for the mediator gluon to decay into a new quark pair. As illustrated in Figure 1.6, when the two charm quarks are pulled apart, the strong force between them increases, and the mediating gluon will have a large amount of energy. Before the $c\bar{c}$ quarks are further separated, the gluon creates a pair of $d\bar{d}$ quarks. This process continues, and eventually the particle formed by $c\bar{c}$ quarks will become two particles formed by $c\bar{d}$ and $\bar{c}d$ quarks. The gluon could create any pair of quarks, as long as its energy is large enough. Here we use $d\bar{d}$ quark pair as an illustration. Also $d\bar{d}$ quark pair requires relatively low energy to create, compared to other quark anti-quark pairs.

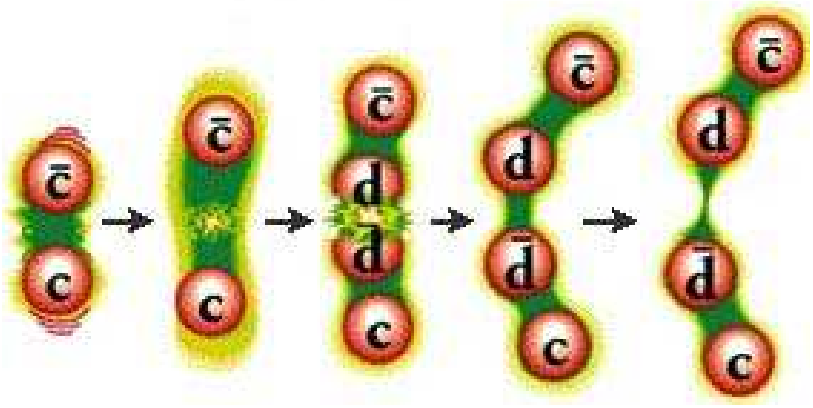

Figure 1.6: The illustration of QCD con nement [8].

Jet formation

When quarks and gluons are created in a high energy collision, they will move away from the collision position freely for a brief moment. And then, because of QCD con nement, when the quarks are separated by a distance $\gg 1\ fm$, new quark pairs are created from the virtual gluons exchanged by the interaction between the initial quarks, as shown in Figure 1.7. This process stops when the gluons or quarks don t have enough kinetic energy to create new quark-aniquark pairs. The quarks or gluons initially created by the the energetic collision, enentually become hadrons; this process is called hadronization. The stream of particles created by the hadronization of a single quark or gluon is called a jet, as shown in Figure 1.7 and 1.8.

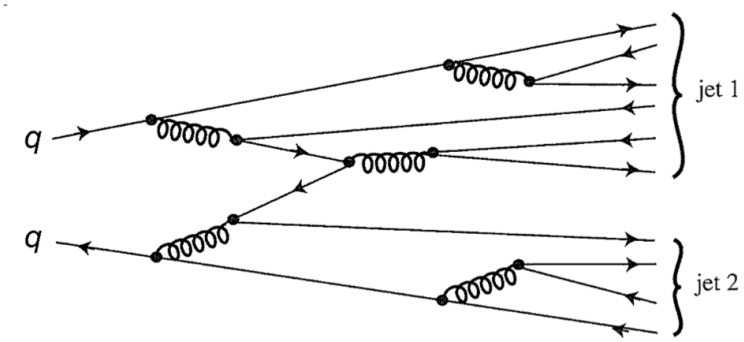

Figure 1.7: The hadronization process [6].

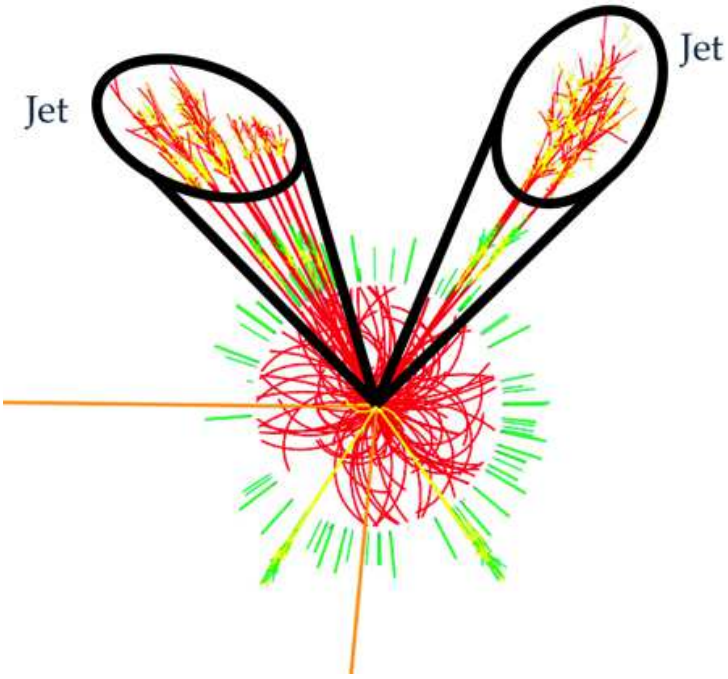

Figure 1.8: A sketch of jets formation in a high energy collision. Hadrons are clustered together to make jets.

1.2.3 The SM Higgs mechanism

The SM Higgs boson is a scaler boson, with spin 0. In SM, Higgs has a non-zero vacuum expectation value (VEV), as shown in Figure 1.9, while all other particles have zero VEVs. As shown in Figure 1.9, the Higgs particle with the non-zero VEV is tending to slide down to the bottom of the potential. While the Higgs particle is on top of the potential, a rotation of the whole system in space-time dimensions, does not change its symmetry. However, when the Higgs particle is sliding off the potential to the bottom, as shown in Figure 1.9, the rotation symmetry is broken. This is called the spontaneous symmetry breaking. The Higgs field with non-zero VEV is permeating all the space. And fermions, by their interactions with the Higgs particle, gain their masses. The magnitude of a fermion's mass is proportional to the its coupling strength with the Higgs field.

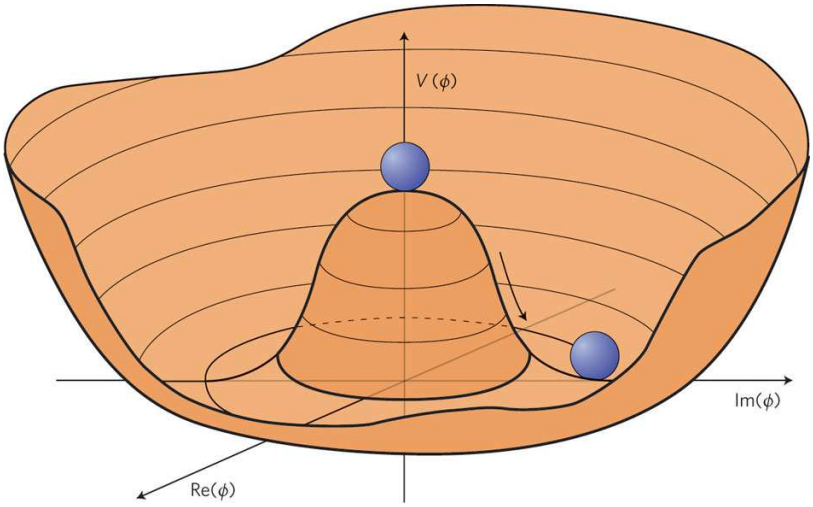

Figure 1.9: The non-zero vacuum potential of Higgs field. The $Im(\phi)$ and $Re(\phi)$ axes represent the plane of space and time. The $V(\phi)$ axis represents the potential energy. And the circle on top of the concave potential is the Higgs particle.

The W^+, W^- and Z bosons gain their masses through the spontaneous electroweak symmetry breaking mechanism. In an unbroken unified electroweak theory, there are four types of massless bosons: W_1, W_2, W_3, and X. And there are also four types of Higgs particles, which can not be distinguished from each other. After spontaneous symmetry breaking, these four Higgs particles become distinguishable: charged H^+, H^-, and neutral H_0 and h. The W_1 boson coupling with the H^+ becomes the massive W^+ boson. The W_2 coupling with the H^- becomes the W^- boson. The W_3 boson combined with the X boson together coupling with the H_0 becomes the Z boson. And the residual component of W_3 and X combination becomes the massless photon.

The h, as one of the four Higgs bosons, is not absorbed by other particles, which is called the Higgs particle in SM. However, there is no constraint on the mass of this Higgs particle in SM.

1.3 The physics beyond the SM

Although the SM explains a lot of facts of current experiments and also achieves another tremendous success on the discovery of the SM-like Higgs boson in 2012. However, the SM cannot explain several important phenomena, and thus it is believed to be an effective theory of a more fundamental theory.

Dark matter and dark energy

As shown in Figure 1.10, the universe is expanding. However, the expansion is

accelerating, instead of slowing down. Thus there must be some mysterious force overcomes the attractive force of gravity ad causes the accelerated expansion with time. This unknown force, is usually referred to as "dark energy". Dark here is the thing invisible to us.

The need for dark matter arises from the astronomy observations that the rotational motion of the stars or galaxies suggests a 5∼10 times larger gravitational force than the one could be provided by the matter of the clusters. The lack of matter in this kind of case indicates the existence of matter that couldn't be observed by us, which is called "dark matter".

The current compositions of matter and energy of our universe, from studies, are about 4.9% normal matter (like stars, planets, etc.), 26.8% dark matter, and 68.3% dark energy. The neutrinos of the SM, are stable and have tiny masses, and also interact weakly with other particles. So on a cosmic scale, they behave like the dark matter. However, the redundancy of dark matter (26.8%) over normal matter (4.9%) suggests that neutrinos are insufficient to explain the observed amount of dark matter.

The hierarchy problem

As shown in Table 1.1, the gravitational force is quite small compared to other three forces. It is about 10^{-30} smaller than the weak force. This large discrepancy of scale is called the hierarchy problem of the SM.

The more formal way to state the hierarchy problem is why the Higgs mass in SM is in the scale of ≈125 GeV, rather not 10^{18} GeV, while the latter is more natural [10].

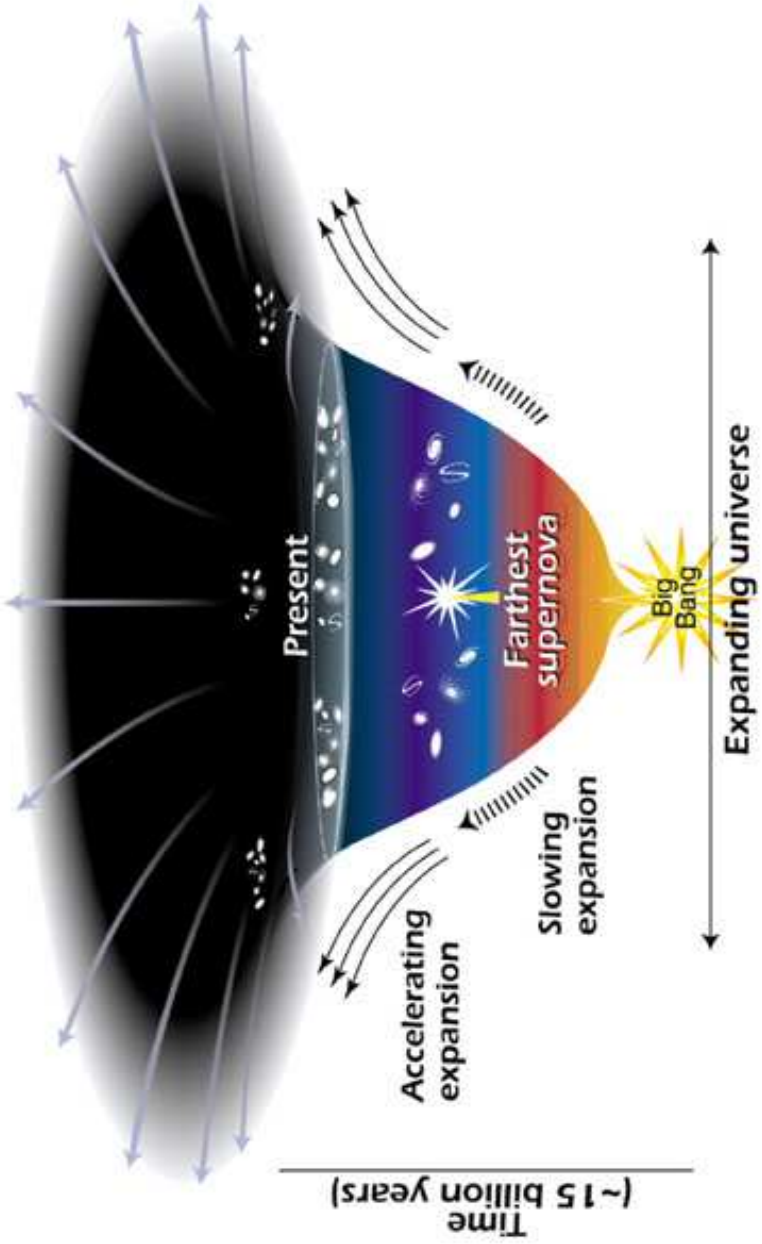

Figure 1.10: Accelerating expansion of the universe [9].

Baryon-anitbaryon asymmetry

The imbalance in baryonic matter and anti-baryonic matter in our observed universe can not be explained by the SM, while the Big Bang should produce equal amount of matter and antimatter.

Gravity

The graviton is the hypothetical mediator of the gravitational force in SM. However, it has not been found yet. Unlike the QED and QCD, there is no known way to include gravity in SM.

Since SM leaves us with a number of mysteries, the research towards a better understanding of our universe has never been stopped. Many theories beyond the SM have been proposed. Composite models of quarks (q*) [11,12] with their potential to explain the generation structure of quarks have been quite popular. The Randall-Sundrum model with its potential to solve the hierarchy problem in SM predicts the existence of Randall-Sundrum Graviton (G_{RS}) [13,14]. There are also extensions of Randall-Sundrum models predicting the existence of bulk Graviton (G_{Bulk}) [15–17]. Many theories beyond the SM also predict the existence of W′and Z′ [18], the heavy partners of the SM W and Z bosons.

According to experimental measurements, these predicted resonances are expected to have resonance masses at least a few hundred GeV. The Large Hadron Collider (LHC), with its high collision energy, is likely to produce these massive resonances.

The predicted massive resonances decaying into a quark and a W or Z vector boson, or into two bosons (WW, ZZ, WZ, WH, or ZH) are searched in this thesis.

In proton-proton (pp) collisions at the energies reached at the LHC, bosons emerging from such decays usually would have sufficiently large momenta so that the hadronization products of their $q\bar{q}'$ decays would merge into a single massive jet [19]. So the event has a dijet topology. In this thesis, two dijet studies for physics beyond the SM are conducted with the CMS detector at LHC for the experimental search of q^*, G_{RS} , G_{Bulk}, W′and Z′.

Chapter 2

The CMS detector at the LHC

The LHC overview

The Large Hadron Collider (LHC), shown in Figure 2.1, is the world′s largest and most powerful particle accelerator. It consists of a ≈27 km ring (underground tunnel) of superconducting magnets, on the border of France and Switzerland. Inside of the ring, two energetic particle beams travel at approximate speed of light before colliding. The two particle beams travel in opposite directions in beam pipes kept in ultrahigh vacuum. A strong magnetic field provided by superconducting magnets guides the two particle beams. The magnets are built from superconducting electric coils kept in a temperature ≈ -271.3°C. To cool the system, much of the accelerator is connected to a distribution system of liquid helium. Thousands of magnets are used in LHC, in which ≈1000 dipole magnets with 15 m in length ensure the circular orbits of the two beams and ≈400 quadrupole magnets, each with 5∼7 m in length,

focus the beams.

At LHC, each proton beam has ≈2000 bunches and each bunch has $\approx 10^{11}$ protons. There is one bunch crossing (collision of bunches) every 25 ns. The large collision energy, the enormous amount of collisions and especially the large rate of bunch crossing raise a big challenge for detecting the events produced at LHC.

Along the beam line of LHC, there are four main detectors: CMS, ATLAS, LHCb and ALICE. The focus of this chapter is to present a brief overview of the Compact Muon Solenoid (CMS) detector. We will start by introducing the coordinate system of CMS.

The coordinate system of CMS

In CMS, the z-axis is the along the beam line. The y-axis is vertically upward and the x-axis is directed radially inward toward the center of the LHC ring. As shown in Figure 2.2, the beam line, which is the z-axis, is perpendicular into this paper. And the x- and y-axes are on this paper but perpendicular to each other.

The default CMS x-y-z coordinate system and also the r-θ-ϕ coordinate system are both right-handed. In the transverse plane (x-y plane), which is shown in Figure 2.2, the azimuthal angle ϕ is defined as the angle measured from the x-axis ($\tan\phi = y/x$). And the transverse momentum p_T is defined as $p_T = \sqrt{p_x^2 + p_y^2}$. The polar angle θ is defined with respect to the the positive z-axis ($\tan\theta = \sqrt{x^2 + y^2}/z$).

The pseudorapidity η, as shown in Figure 2.3 is defined as $\eta = -\ln \tan[\frac{\theta}{2}]$. The rapidity is defined as $y = \frac{1}{2}\ln\frac{E+p_z c}{E-p_z c}$.

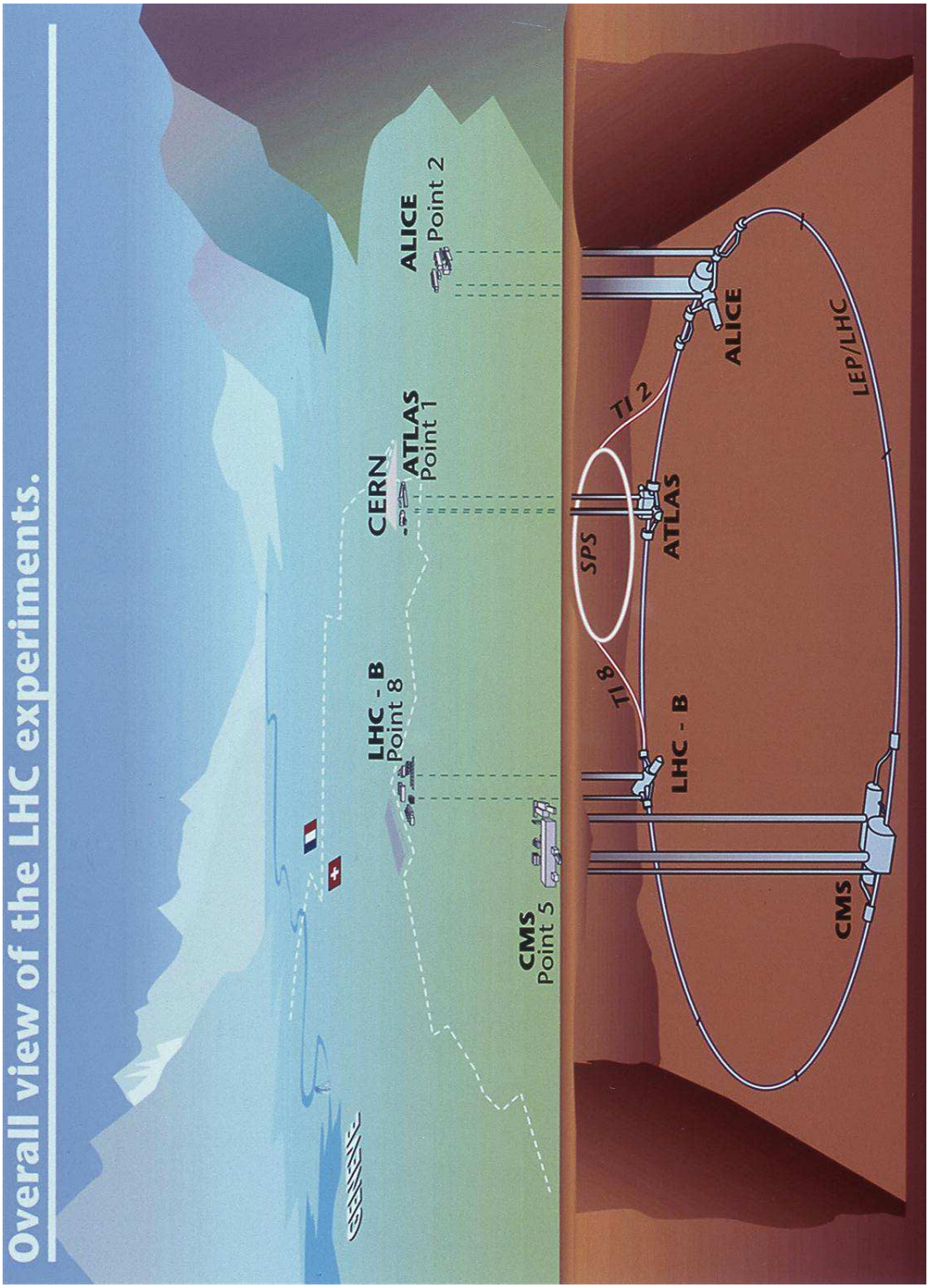

Figure 2.1: An overview of the LHC [20].

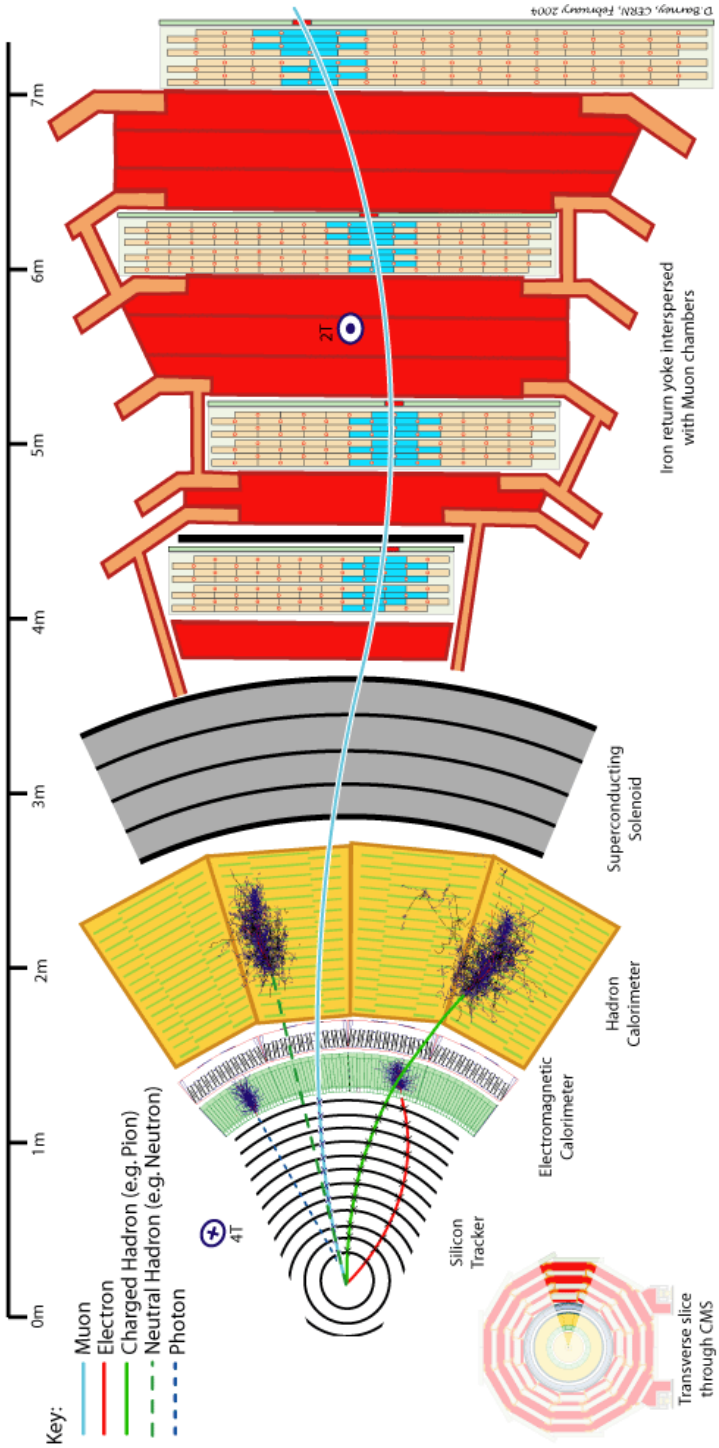

Figure 2.2: Transverse picture of the CMS detector.

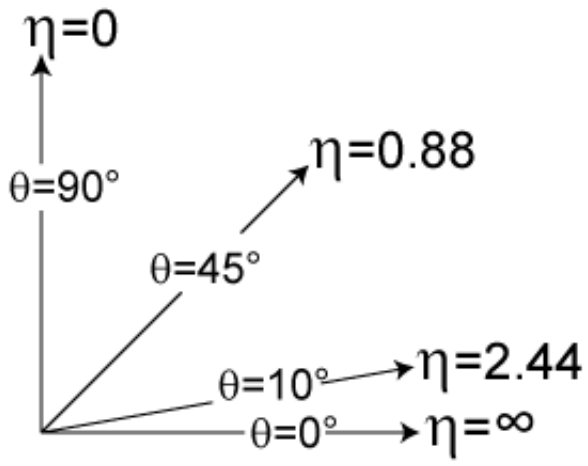

Figure 2.3: The mapping between the pseudorapidity and the polar angle ϕ [21].

The CMS detector overview

CMS is short for Compact Muon Solenoid, which indicates its profession in muon detecting. The overall layout of CMS from di erent viewpoints are shown in Figure 2.2 and Figure 2.4. In Figure 2.2, the z axis is perpendicular into this paper. In Figure 2.4, the z axis is on this paper, though the center of the detector.

The dimensions of the CMS detectors are a length of 21.6 m, a diameter of 14.6 m and a total weight of 12500 tons. From the beam line to the outside exterior, there are silicon tracker, Electromagnetic Calorimeter, Hadronic Calorimeter, Superconducting Solenoid, and Muon stations.

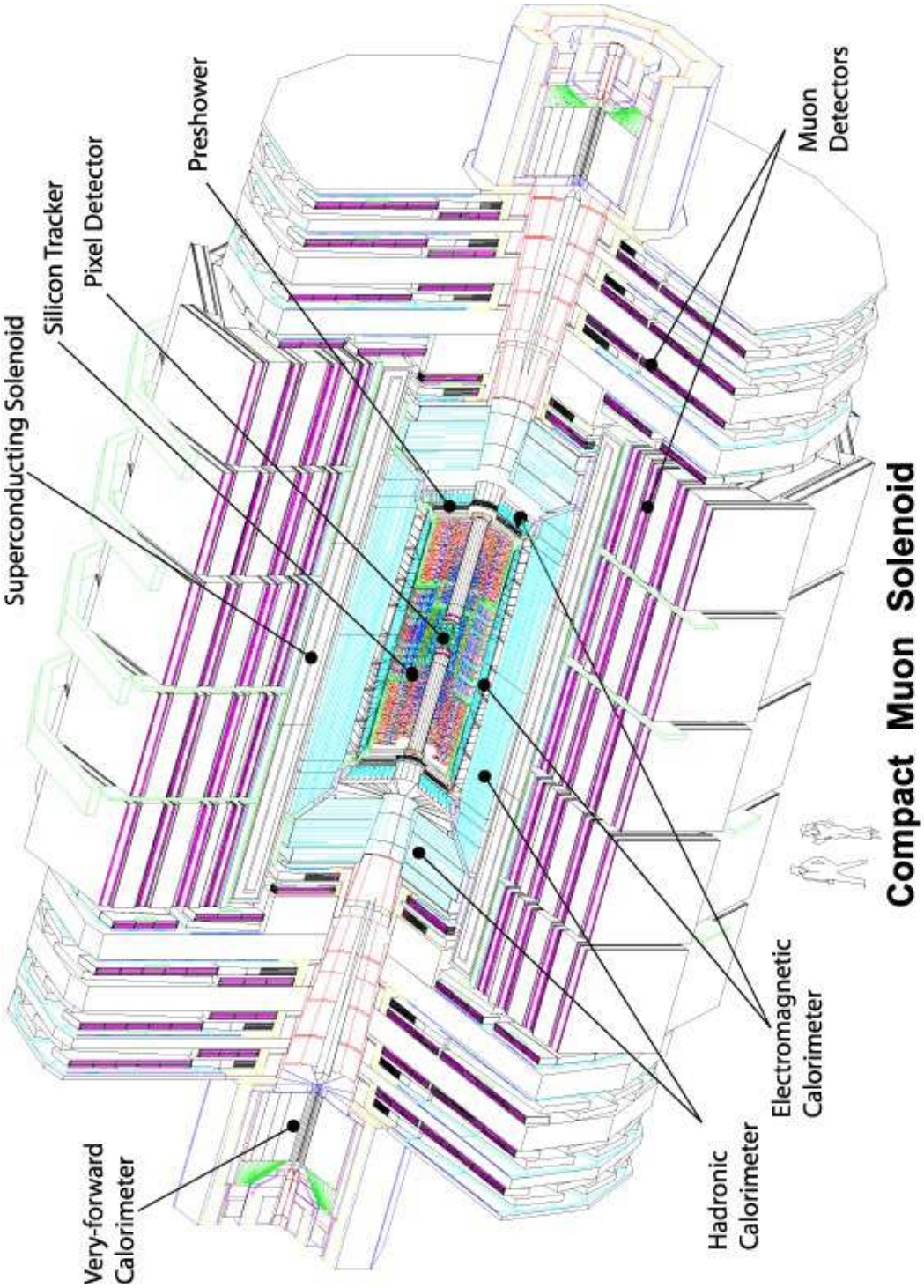

Figure 2.4: Overview of the CMS detector along the beam line.

2.1 The Magnet

Each proton beam of the LHC had energy of 4 TeV in 2013, and it will reach 6.5 TeV in 2015, and 7 TeV in 2016. Particles of interest from such energetic collisions are expected and likely to have high p_T. So in CMS, to achieve a better momentum resolution of the high-p_T charged particles, a large magnetic field of 3.8 T is chosen. For a current of 1 A in a loop of radius 3 m, the resulting magnetic field is $\approx 10^{-7}$ T. So in CMS, to generate a field of 3.8 T for a radius of $\approx$3 m, a large current 19.5 kA is applied, with 2168 turns of coil. CMS solenoid uses a high-purity aluminium-stabilised conductor and indirect cooling by thermosyphon to achieve superconducting.

The radius of the CMS solenoid is chosen to be large enough to accommodate both the inner tracker and the calorimetry inside. The detailed parameters of this setup are shown in Table 2.1.

2.2 The inner tracking system

As shown in Figure 2.2, the particles resulting from the collision point first pass through the silicon tracking detector. The large rate of collisions will produce a substantial flux of particles. For the tracker system, beside the challenge of radiation damage, the response time must be very small and the sensors need to have good spacial resolution. These two aspects are the main design targets of the silicon tracker system.

Table 2.1: Parameters of the CMS superconducting solenoid.

Characteristics	Values
Field	4 T
Inner Bore	5.9 m
Length	12.9 m
Number of turns	2168
Current	19.5 kA
Stored energy	2.7 GJ
Hoop stress	64 atm

The overall tracking volume is given by a cylinder of length of 5.8 m and diameter of 2.6 m. It is mainly composed of two parts: the inner pixel tracker and the layers of the outer strip tracker. The schematic diagram of the 1/4 of the CMS tracker system (along the z axis) is shown in Figure 2.5.

Three layers of silicon pixel detectors are placed closed to the interaction region to improve the measurement of the impact parameter [1] of charged-particle tracks, as well as the position of secondary vertices. In addition, CMS uses 10 layers of silicon microstrip detector, which provide the required granularity and precision.

The pixel detector

[1] The impact parameter is defined as the perpendicular distance between the path of a projectile and the center of the field created by an object that the projectile is approaching. http://en.wikibooks.org/wiki/LaTeX/Footnotes_and_Margin_Notes

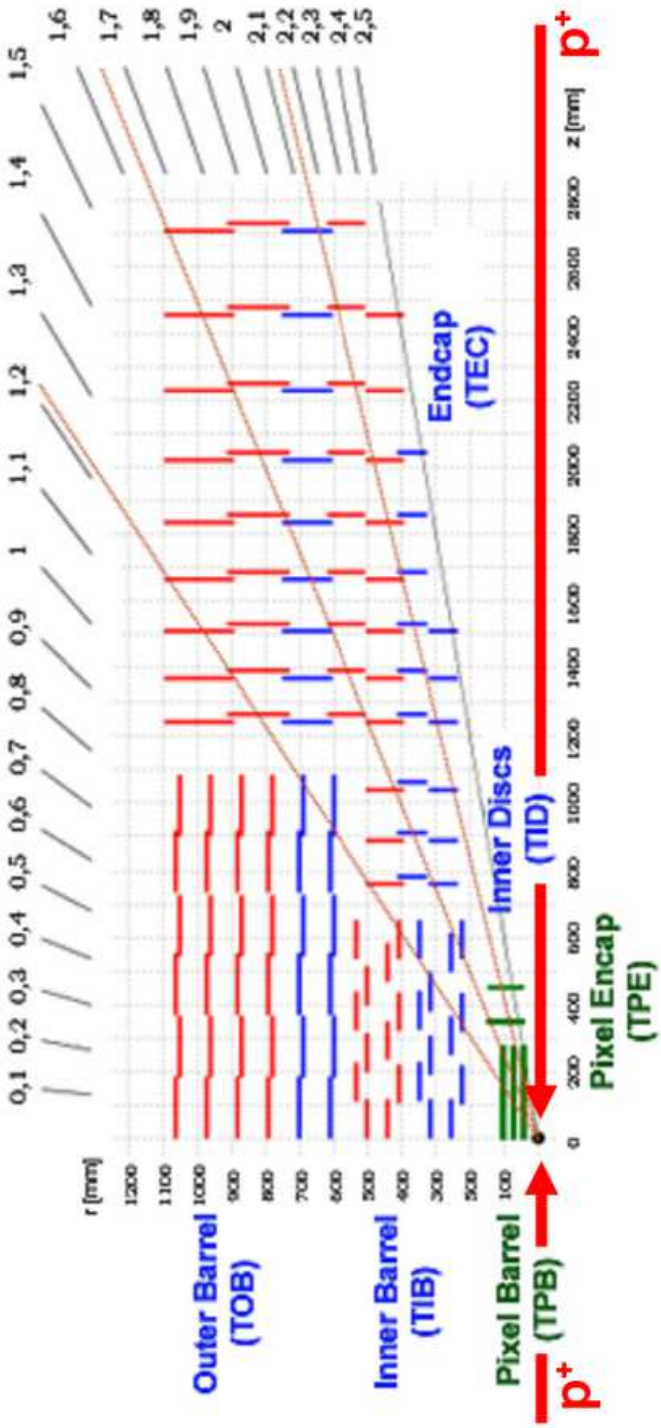

Figure 2.5: The tracker layout of CMS system.

Close to the interaction vertex, in the barrel region, are 3 layers of hybrid pixel detectors at a radii of 4.4, 7.3, and 10.2 cm, as shown in Figure 2.6. Each layer is spilt into sensor segments like mosaic tiles. Each silicon sensor, with the size of 100 150 (m)^2, is about two hairs widths. The endcap of pixel detector is composed of two disks of pixel modules on each side of the barrel region, extending from 6 to 15 cm in radius.

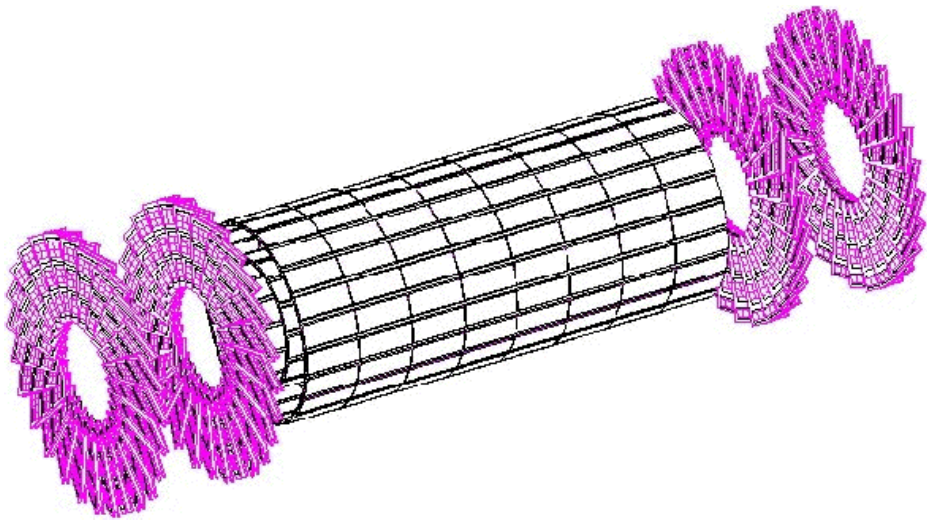

Figure 2.6: Layout of the pixel detector in the CMS [22].

When a charge particle passes trough the sensor, it raises an electronic signal. Knowing which pixels have collected charges allows us to reconstruct the charged particle s trajectory. Because the pixel detector is made of 2D tiles, and has three layers, a three-dimensional picture of the particle s motion is created. The spatial resolution of the pixel detector is 10 m for the r measurement and 20 m for the z measurement.

The strip detector

After passing through the three pixel layers, particles travel through ten layers of silicon strip detectors. The strip system is shown in Figure 2.7, and reaches an outer radius of 130 cm. The silicon strip tracking detector consists of four inner barrel (TIB) layers assembled in shells with two inner endcaps (TID), each composed of three small discs. The outer barrel (TOB) consists of six concentric layers. Finally two endcaps (TEC) close off the tracker. Each part of the strip detector has silicon modules designed differently for its place within the detector.

Unlike the 2D pixel sensor, most of the 10 strip layers are composed of 1D strip sensors. So when a charged particle passes through the strip sensor, the strip detector only outputs the local 1D position instead of 2D position.

2.3 Electromagnetic calorimeter (ECAL)

The ECAL is designed to calibrate the energy of electron and photons resulting from proton-proton (pp) collisions at LHC, and its structure is shown in Figure 2.8. ECAL uses lead tungstate ($PbWO_4$) crystals with coverage in $|\eta|$ up to 3.0. The lead tungstate crystal is highly transparent and "scintillates" when electrons and photons pass through it, which produces light in proportion to the charged particles′ energy. It also has short radiation (0.89 cm) and Moliere (2.2 cm) lengths, which allows the ECAL to be hermetic for electrons and photons and also more compact. It is also

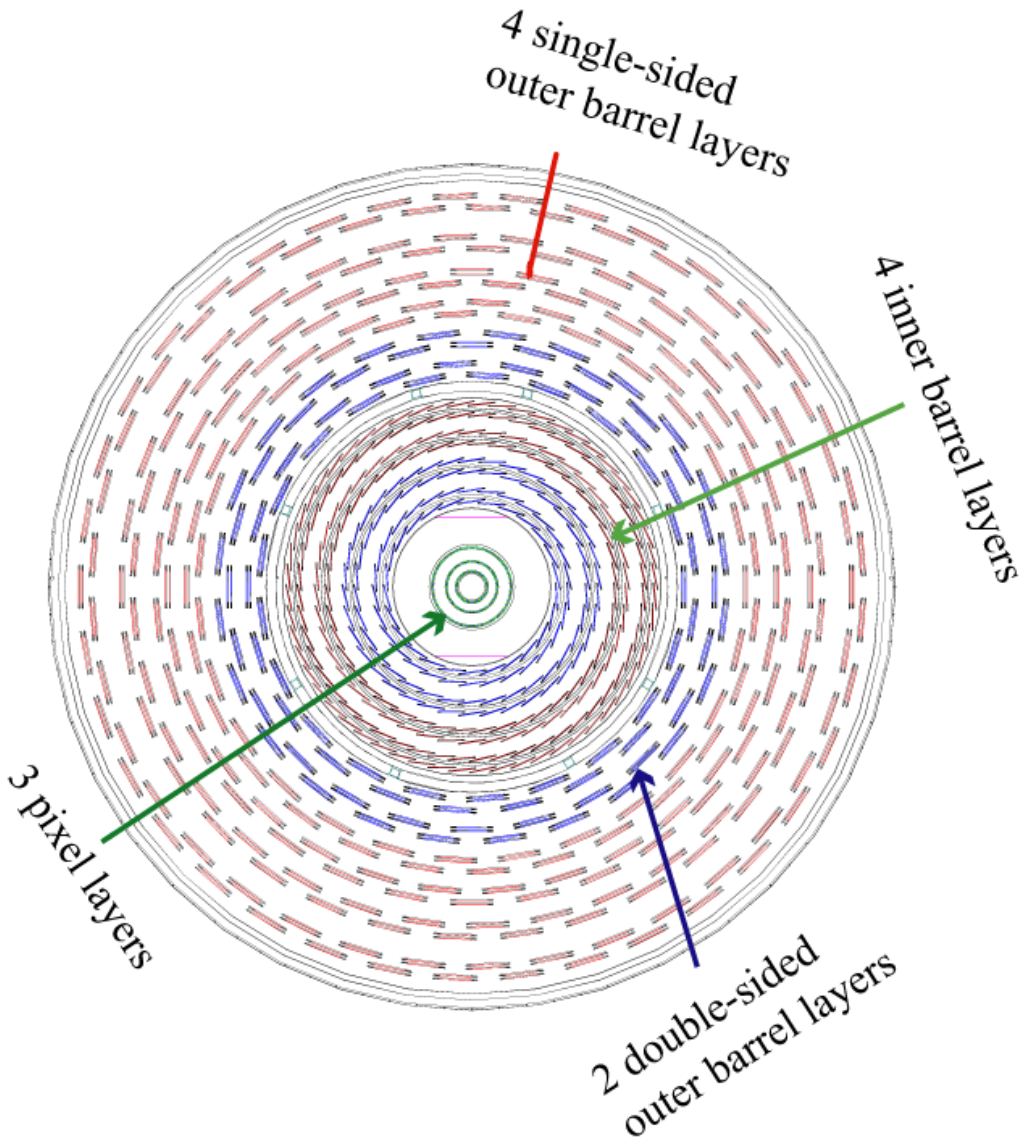

Figure 2.7: Layout of tracker system of CMS, with z-axis perpendicular into the paper.

fast and radiation hard. However, the $PbWO_4$ crystal produces relatively low light yield. So the silicon avalanche photodiodes (APDs), which could convert the light signal into amplified electric signal, are used as photodetectors in the ECAL barrel (EB) and vacuum phototriodes (VPTs) in the ECAL endcap (EE).

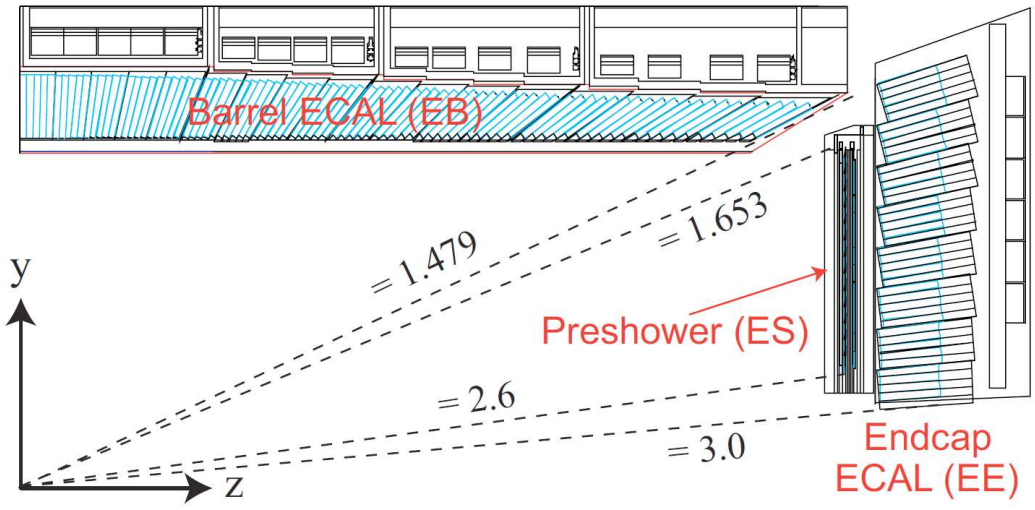

Figure 2.8: Geometric view of one quarter of the ECAL.

These photodetectors have been especially designed to work within the high magnetic field. They are glued onto the back of each of the crystals to detect the scintillation light and convert it to an electrical signal which is then amplified and sent to the data acquisition system.

A preshower system is installed in front of the EE. The preshower is made of two planes of lead followed by silicon sensors. The reason for the preshower system is that short-lived particles called neutral pions, produced in pp collisions, can inadvertently mimic high-energy photons when they decay into two closely-spaced lower energy

photons that the ECAL picks up together. One of the main channels for SM Higgs discovery is the signature from $H \to \gamma\gamma$, in which the final state γ particles are likely to have a high energy at LHC. So the preshower system could identify the photons from neutral pion decay and distinguish them from the photons of $H \to \gamma\gamma$ decay.

2.4 Hadronic calorimeter (HCAL)

HCAL is designed to measure the energy of hadrons and also the transverse missing energy E_T^{miss} (MET), the imbalance in the transverse momentum of all visible particles. MET can be used to study not only the neutrino (invisible) related processes in SM, but also many models beyond the SM predicting new invisible particles, e.g., Dark Matter models, supersymmetric models, and models with large extra dimensions. Thus improving the energy resolution and achieving good hermeticity for the the E_T^{miss} measurement, are the two main goals of HCAL design.

As shown in Figure 2.9, HCAL is composed by four parts, HCAL barrel (HB), HCAL endcap (HE), HCAL outer (HO) and HCAL forward (HF). HF, not presented in the plot, sits on the outside of the muon stations and covers $3 < |\eta| < 5.0$.

Brass is the filling material of HCAL. It is chosen because it is non-magnetic and has short interaction length. To achieve a good containment, HCAL maximizes the brass inside of the solenoid by minimizing the detection material with the application of the tile/fiber technology [22]. HCAL is a sample detector, in which layers of brass

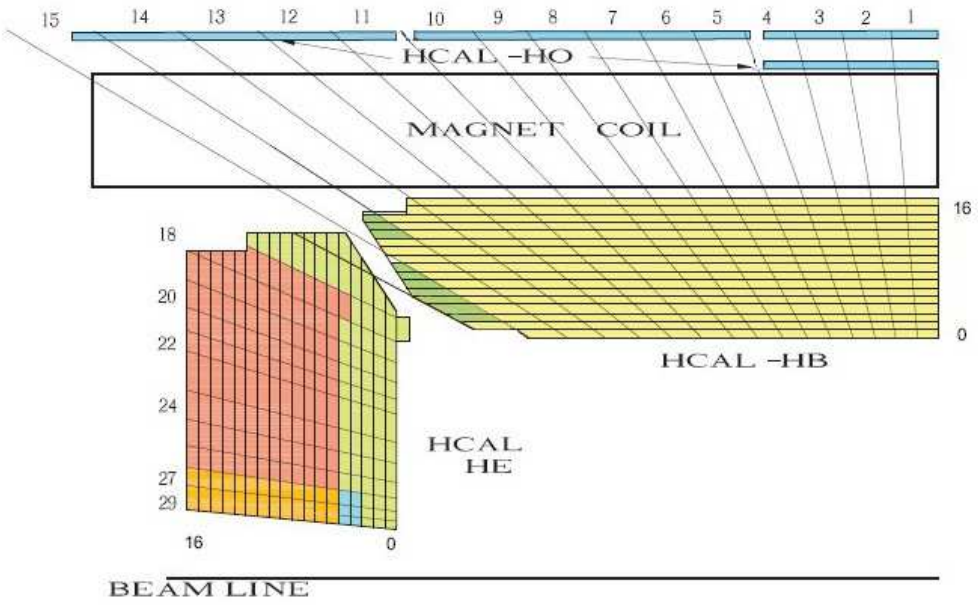

Figure 2.9: Geometric view of one quarter of the HCAL.

are interleaved by layers of fiber. The HO is located outside of the solenoid, to complement the measurement of HB.

2.5 Muon system

The layout of one quarter of the CMS muon system for the initial low luminosity running is shown in Figure 2.10, and the transverse view of the muon stations (MSs) is shown in Figure 2.11, with z-axis perpendicular into the paper. In Figure 2.11, the red colored part is the return yoke, which has a magnetic field of 2 T. As shown in Figures 2.10 and 2.11, in the Muon Barrel (MB) region, 4 stations of detectors

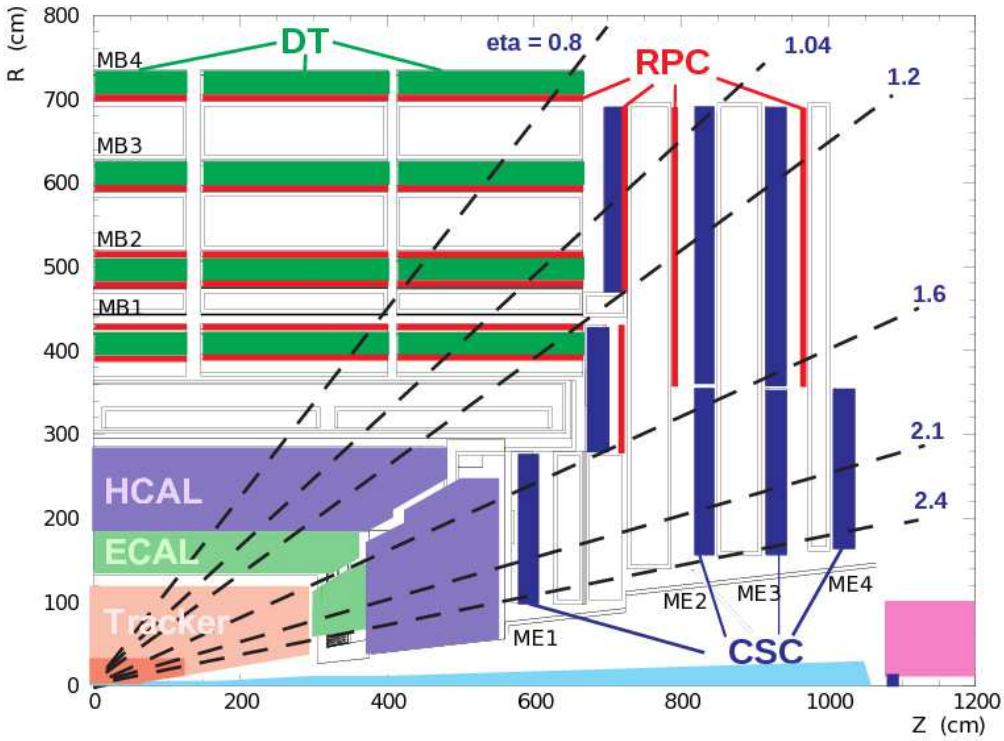

Figure 2.10: Layout of one quarter of the CMS muon system for initial low luminosity running. The RPC system is limited to $|\eta| < 1.6$ in the endcap, and for the CSC system only the inner ring of the ME4 chambers have been deployed.

are arranged in cylinders interleaved with the iron yoke. This magnetic field in the return yoke bends the trajectory of muon, while there is almost no magnetic field in the four muons stations (MS1, MS2, MS3, MS4). In the adjacent muon stations by comparing the bending angle because of the return yoke , the muon system correctly calculates the muon momentum.

From Figure 2.10, three types of gaseous detectors are used to identify and measure muons: drift tube (DT), cathode strip chamber (CSC), and resistive plate chamber (RPC).

Drift tube (DT), with detailed layout in Figure 2.12, is used in the barrel region (

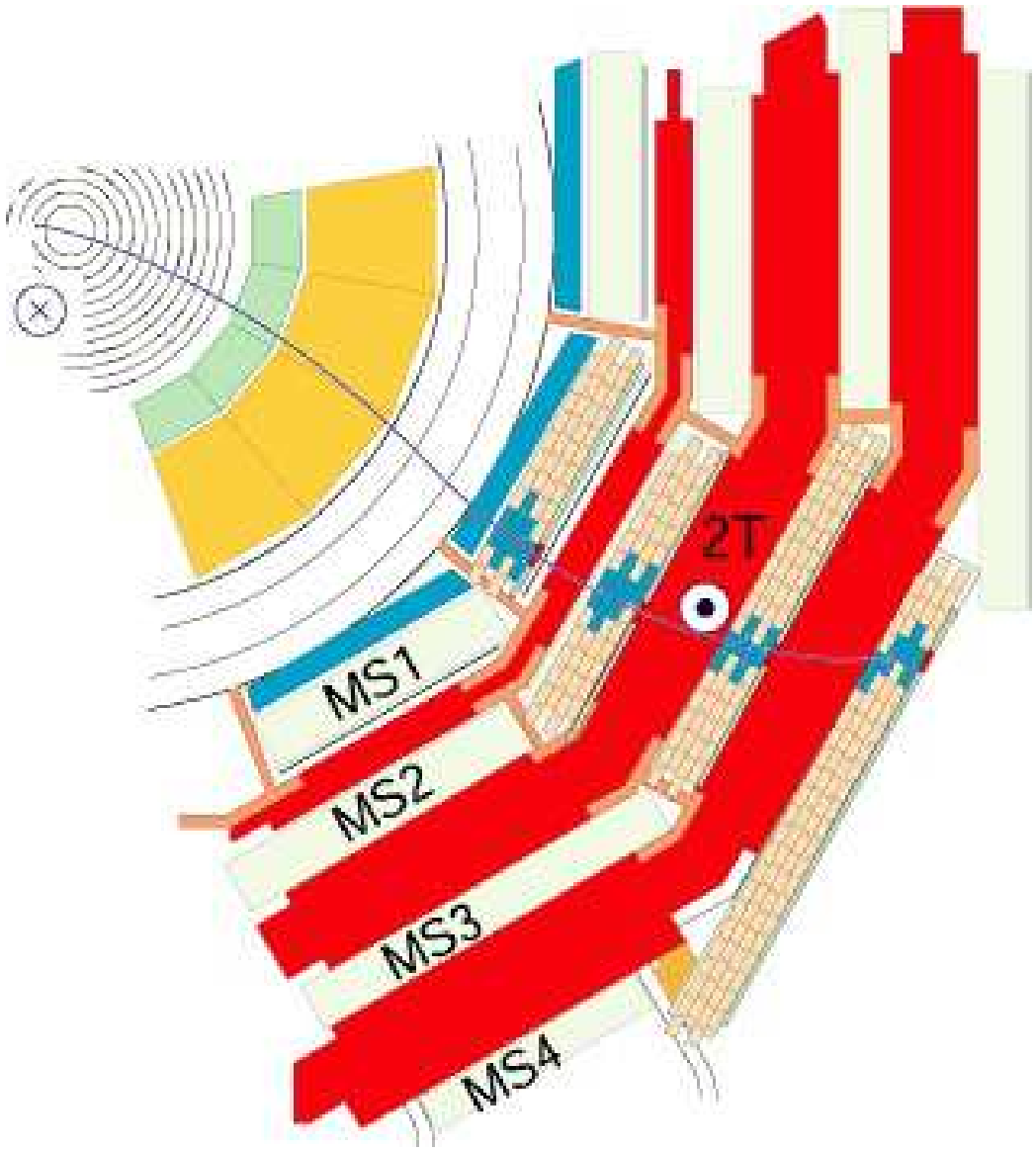

Figure 2.11: The Muon stations in the transverse view, with z-axis perpendicular into the page.

$|\eta| < 1.2$). In this region, the residual magnetic field in the chambers is low and also muon rate is low, so drift tube is chosen. When a muon or any charged particle passes through the gas volume, it ionizes the atoms of the gas. The free electrons follow the electric field ending up at the positively-charged wire (anode wire in Figure 2.12). By registering the location on the wire and also the time the electrons take to reach the wire, the DT could provide a 2D position of the passing charged particle. The maximum drift length is 2.0 cm and the single point resolution is $\approx 200\ \mu m$.

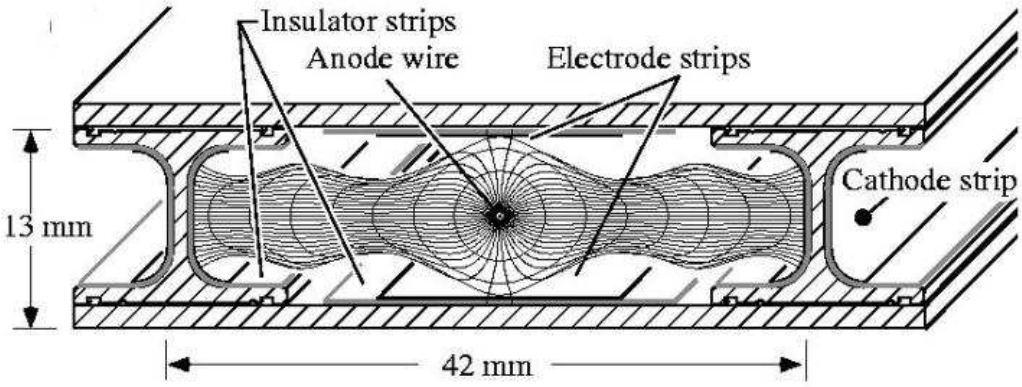

Figure 2.12: Layout of the drift tube [23].

In the endcap region, cathode strip chambers (CSCs) are used because of the high residual magnetic field and also high muon rate. As shown in Figure 2.13, CSCs

consist of arrays of positively-charged "anode" wires crossed with negatively-charged copper "cathode" strips within a gas volume. When muons pass through, they knock electrons off the gas atoms, which flock to the anode wires producing an avalanche of electrons. Positive ions move away from the wire and towards the copper cathode, also inducing a charge pulse in the strips, at the right angles to the wire direction. Because the strips and the wires are perpendicular, we get two position coordinates for each passing charged particle.

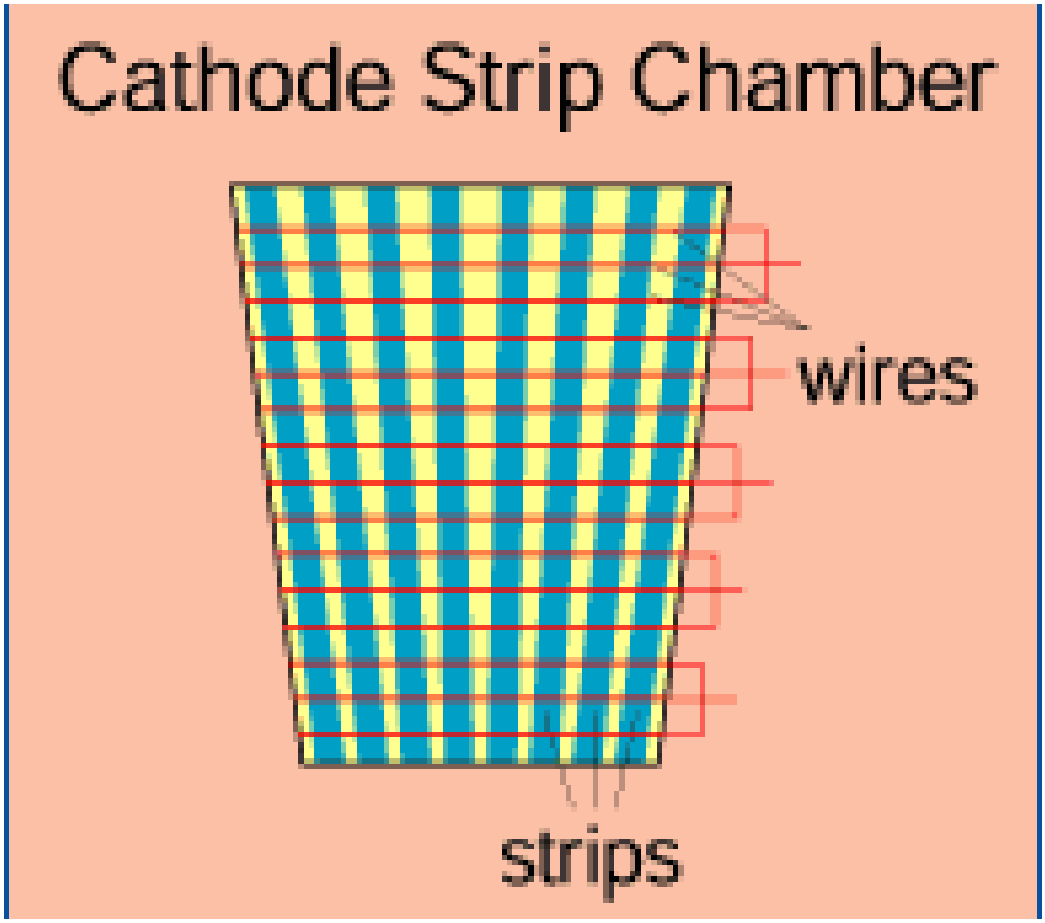

Figure 2.13: Layout of the drift tube [24].

In addition to this, resistive plate chambers (RPCs), as shown in Figure 2.14, are used in both the barrel and the endcap regions. The RPCs could provide a fast response with a good time resolution but with a coarser position resolution than the DTs and the CSCs. So the RPCs could identify the correct bunch crossing (25 ns per bunch crossing).

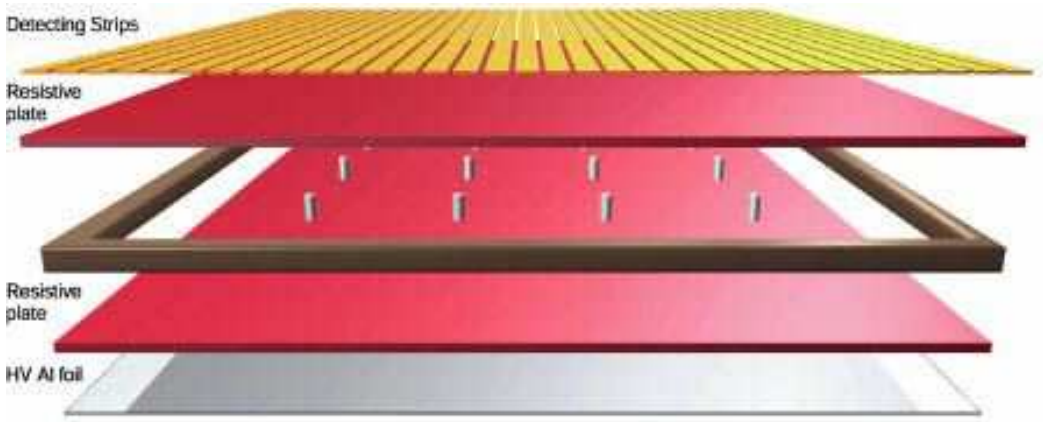

Figure 2.14: Layout of the resistive plate chamber [24].

Muons from pp collisions are measured 3 times: in the silicon tracking system, after the solenoid coil, and in the muon chambers (the muon system). Measurement of the momentum of muons using only the muon system, is essentially determined by the muon bending angle when it exits the 4 T solenoid, taking the interaction point of pp collision as the origin of the muon. For low-momentum muons, the best momentum resolution is given by the resolution obtained in the silicon tracker. For high-momentum muons, combining the inner tracker and muon detector measurements significantly improves the muon momentum resolution. At CMS, in $0 < |\eta| < 2.0$, for

μ with p_T 200$\sim$400 GeV, the relative error on muon's momentum $\Delta p/p$ is measured to be $\leq 3\%$.

2.6 Trigger and data acquisition

As mentioned earlier, there is one bunch crossing per 25 ns. While for each bunch crossing, there is $\approx$40 millions pp collisions. Since we are only interested in events that further our understanding of the SM and possibly indicate new physics, the CMS adopted a two-level trigger system to filter out the uninteresting events.

The Level-1 (L1) trigger is automatic and universally applied to each event, with the application of hardware processors. Basically, L1 trigger sets a threshold for "trigger primitive" objects, like photon, muon, electron, and jets to be above some E_T or p_T. After the L1 trigger, the event rate reduces to 100 kHz.

The L2 trigger, which is also named as high level trigger (HLT), is used to further reduce the 100 kHz to 100 Hz event rate. In L2 trigger system, the event from pp collision is partially reconstructed. Information about the calorimeter and muons is first reconstructed and compared with the threshold of L2. Events falling this threshold will be immediately thrown out. Then pixel tracks are reconstructed and tested with the corresponding threshold. Following this kind of process, without reconstructing all possible objects in an event, L2 is more flexible and has complete freedom in selecting events.

Events passing all levels of the trigger system, are fully reconstructed and written to the disks.

Chapter 3

Search for $X \to qV$ *or* VV at LHC at $\sqrt{s} = 8$ TeV

3.1 Introduction

As we mentioned in Chapter 1, the SM is limited and couldn't provide solutions for some important phenomena. Several models of physics beyond the standard model (SM) predict the existence of resonances with masses above 1 TeV that decay into a quark and a W or Z vector boson, or into two vector bosons. In proton-proton (pp) collisions at the energies reached at the Large Hadron Collider (LHC), vector bosons emerging from such decays usually would have sufficiently large momenta so that the hadronization products of their $q\bar{q}'$ decays would merge into a single massive jet [19]. We present a search for events containing one or two jets of this kind in pp collisions

at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The data sample, corresponding to an integrated luminosity of 19.7 fb^{-1}, was collected with the CMS detector at the LHC.

The signal is characterized by a peak in the dijet invariant mass distribution m_{jj} over a continuous background from SM processes, comprised mainly of multijet events from quantum chromodynamic (QCD) processes. The sensitivity to jets from W or Z bosons is enhanced through the use of jet-substructure techniques that help differentiate such jets from remnants of quarks and gluons [25, 26], providing the possibility of "W/Z-tagging". This search is an update of a previous CMS study [27] performed using data from pp collisions at $\sqrt{s} = 7$ TeV. Besides increased data-sample size and larger signal cross sections from the increase in centre-of-mass energy, this analysis also benefits from an improved W/Z-tagger based on "N-subjettiness" variables, introduced in Ref. [28] and defined in Section 3.3.1.

We consider four reference processes that yield one W/Z-tagged or two W/Z-tagged all-jet events: (i) an excited quark q* [11, 12] that decays into a quark and either a W or a Z boson, (ii) a Randall–Sundrum (RS) graviton G_{RS} that decays into WW or ZZ bosons [13, 14], (iii) a "bulk" graviton G_{Bulk} that decays into WW or ZZ [15–17], and (iv) a heavy partner of the SM W boson (W′) that decays into WZ [18].

Results from previous searches for these signal models include limits placed on the production of q* at the LHC as dijet [29–31] or γ+jet [32] resonances, with a q* lighter than ≈3.5 TeV at a confidence level (CL) of 95% [29]. Specific searches for

resonant qW and qZ final states at the Tevatron [33,34] exclude q* decays into qW or qZ with $m_{q*} < 0.54$ TeV, and results from the LHC [27,35] exclude q* decays into qW or qZ for $m_{q*} < 2.4$ TeV and $m_{q*} < 2.2$ TeV, respectively.

Resonances in final states containing candidates for WW or ZZ systems have also been sought [36–39], with lower limits set on the masses of $G_{\rm RS}$ and $G_{\rm Bulk}$ as a function of the coupling parameter $k/\overline{M}_{\rm Pl}$, where k reflects the curvature of the warped space, and $\overline{M}_{\rm Pl}$ is the reduced Planck mass ($\overline{M}_{\rm Pl} \equiv M_{\rm Pl}/\sqrt{8\pi}$) [13,14]. The bulk graviton model is an extension of the original RS model that addresses the flavour structure of the SM through localization of fermions in the warped extra dimension. The experimental signatures of the $G_{\rm RS}$ and $G_{\rm Bulk}$ models differ in that $G_{\rm Bulk}$ favours the production of gravitons through gluon fusion, with a subsequent decay into vector bosons, rather than production and decay through fermions or photons, as the coupling to these is highly suppressed. As a consequence, $G_{\rm Bulk}$ preferentially produces W and Z bosons that are longitudinally polarized, while $G_{\rm RS}$ favours the production of transversely polarized W or Z bosons. In this study, we use an improved calculation of the $G_{\rm Bulk}$ production cross section [15] that predicts a factor of four smaller yield than assumed in previous studies [36,37].

The most stringent limits on W′ boson production are those reported for searches in leptonic final states [40,41], with the current limit specified by $m_{\rm W'} > 2.9$ TeV. Depending on the chirality of the W′ couplings, this limit could change by ≈0.1 TeV. Searches for W′ in the WZ channel have also been reported [37,42,43] and set a lower

limit of $m_{\mathrm{W}'} > 1.1$ TeV.

The data, and the event simulations are described briefly in Section 3.4. Event reconstruction, including details of W/Z-tagging, and selection criteria are discussed in Section 3.2 and 3.3. The systematic uncertainties are discussed in Section 3.8. And Section 3.8.1 presents studies of dijet mass spectra, including SM background estimates. The interpretation of the results in terms of the benchmark signal models is presented in Section 3.10, and the results are summarized in Section 5.

3.2 Dijet analysis with jet substructure tagging

3.2.1 Event display

The event detected by the CMS detector is shown in Figure 3.1 and Figure 3.2. In Figure 3.1, the top image is showing the event in the transverse plane, which is global $\theta - \phi$ axes. The bottom image is showing this event in the $\theta - z$ plane. In Figure 3.2, the top image is showing the global view of this event. And the bottom image is showing the lego plot of the two jets, each of which is composed by two subjets, a term that we will introduce in the following text. More plots of the events of interest are presented in Appendix A.2.

3.2.2 Jet reconstruction

Jets are reconstructed by clustering particles obtained using the particle flow (PF) algorithm [44–46]. The PF procedure identifies each individual particle (a PF candidate) through an optimized combination of all subdetector information. The energy of photons is obtained directly from the ECAL measurement, corrected for suppression effects of energies from calorimetric channels with small signals (referred to as zero-suppression) [47]. The energy of an electron is determined from a combination of the track momentum at the main interaction vertex, the corresponding ECAL cluster

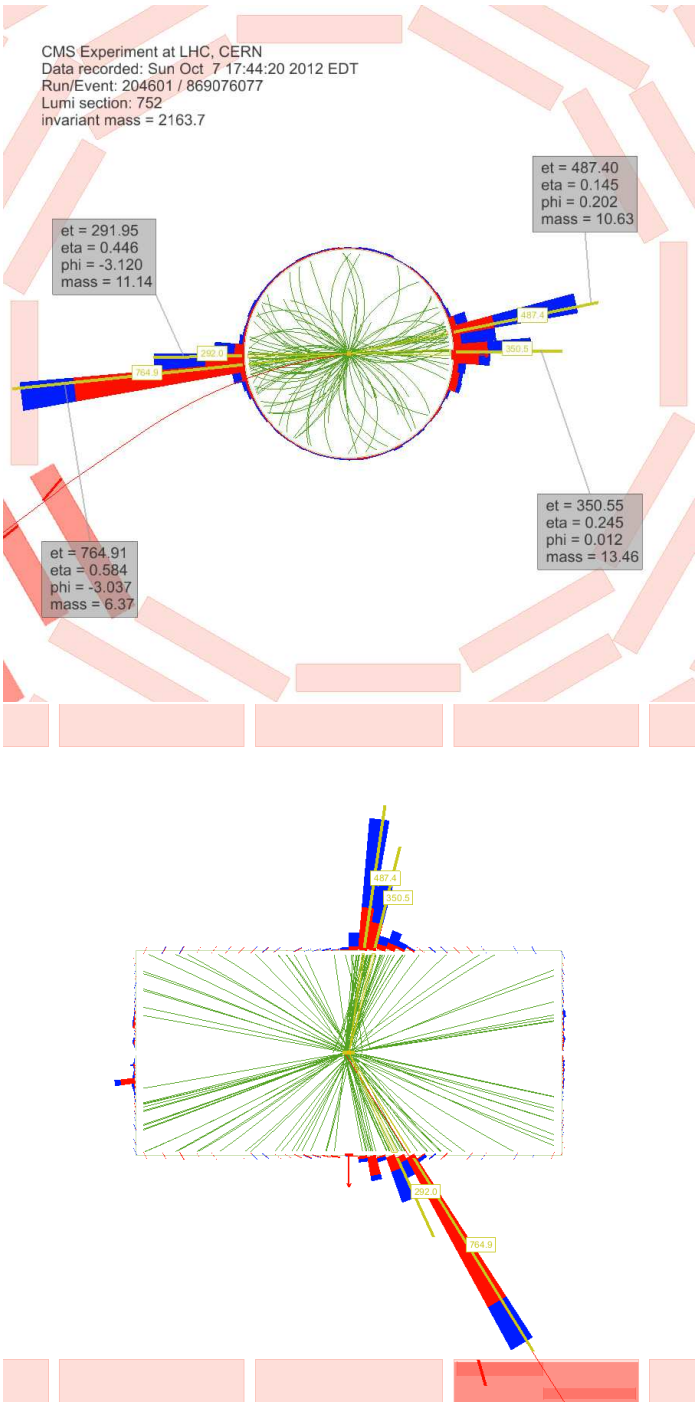

Figure 3.1: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

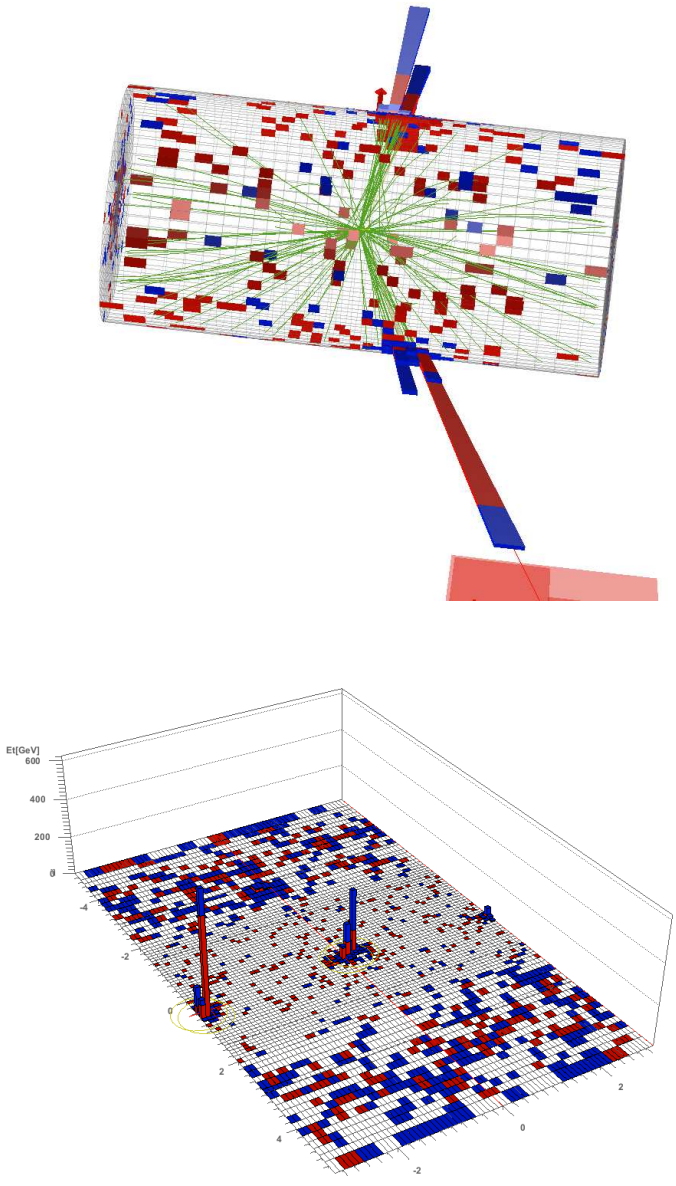

Figure 3.2: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

energy, and the energy sum of all bremsstrahlung photons associated with the track. The energy of a muon is obtained from the corresponding track momentum. The energy of a charged hadron is determined from a combination of the track momentum and the corresponding ECAL and HCAL energies, corrected for zero-suppression effects, and calibrated for the nonlinear response of the calorimeters. Finally, the energy of a neutral hadron is obtained from the calibrated energies in ECAL and HCAL.

The resulting particle flow candidates are passed to each jet clustering algorithm, in this case the Cambridge-Aachen (CA) [48, 49] jet clustering algorithm, as implemented in FastJet version 3.0.1 [50, 51], to create "particle flow jets". The CA clustering sequence is only determined by the distance between clusters and is not weighted by their momentum, as is done for the k_{T} and anti-k_{T} algorithms. A distance parameter of size $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.8$ is used for the CA algorithm. The subjets of the CA8 jets are obtained by rewinding the last step of jet clustering.

Charged hadrons identified as pileup are removed from the inputs to the jet clustering algorithms. The remaining neutral component of pileup is removed by applying a residual area-based correction as described in Ref. [52,53]. The mean $\mathrm{p_T}$ per unit area is computed with the k_{T} algorithm with the "active area" method, with a distance parameter of 0.6, and the jet energy is corrected by the amount of pileup expected in the jet area. The amount of energy expected from the underlying event is added back into the jet. The pileup-subtracted jet four momenta are finally corrected for nonlinearities in η and $\mathrm{p_T}$ with simulated data, with a residual η-dependent correction

added to correct for the difference in simulated and true responses [54, 55].

The jet energy corrections for the CA $R = 0.8$ jets are derived from studies using the anti-k_{T} $R = 0.7$ jet algorithm. Simulation studies confirm that these anti-k_{T}-derived jet corrections are adequate for the CA $R = 0.8$ jet algorithm for the jet momenta considered here [25].

3.2.3 Event selection

Events are selected using the following cuts:

- The event must have a well reconstructed primary vertex as computed by a deterministic annealing filter (DAF) ($|z_{\text{Primary Vertex}}| < 24$ cm, $N_{\text{DOF}} > 6$).
- The following recommended noise event filters are used:
  - CSC tight beam halo filter
  - HBHE noise filter with isolated noise rejection
  - HCAL laser event filter (HBHE) and HCAL laser event filter 2012
  - ECAL dead cell trigger primitive (TP) filter
  - The beam scraping filter
  - Bad EE supercrystal filter
  - The tracking failure filter
  - Good primary vertex filter

  - Tracking coherent noise filter
  - Tracking TOBTEC fakes filter

- The events are required to have at least two ungroomed CA8 jets with
  - $p_T > 30$ GeV , $|\eta| < 2.5$
  - to have muon energy fraction < 0.8
  - pass tight particle flow jet ID. The tight PF jet ID is listed below:
    * Neutral Hadron (EM) Fraction $< 0.90 (< 0.90)$, for all jet η
    * Number of Constituents > 1, for all jet η
    * Charged Hadron (EM) Fraction $> 0 (< 0.99)$, for jet $|\eta| < 2.4$
    * Charged Multiplicity> 0, for jet $|\eta| < 2.4$

- Beam background events are removed using the following requirements:
  - In events with at least 10 tracks, a minimum of 25% of these tracks must be high purity tracks.

- We also require $E_T^{miss} / \sum E_T < 0.5$ to further suppress the noise producing large fake E_T^{miss}.

- The events must pass $|\Delta\eta| < 1.3$, $m_{jj} > 890$ GeV

This sample of dijet events is then tested for presence of hadronically decaying W or Z bosons.

3.3 The W/Z-Tagging algorithm

The products of hadronic decays of W/Z bosons can fall within a single jet if these particles are boosted relative to their mass. The W/Z tagging algorithm is developed to identify these boosted W/Z jets, based on the removal of the soft components of the jets (jet pruning) [25,56].

Jet pruning is implemented as application of additional cuts in the process of CA jet clustering. This algorithm starts from a set of "protojets" given by the PF particles that form the original CA jet within a cone of R = 0.8. As in the standard CA jet clustering, these protojets are iteratively combined with each other until all jets is found; however, here the large angle and low p_T protojets are removed in the process. The same parameters are chosen for the jet pruning algorithm as in the original theoretical papers [57,58].

Besides the jet pruning, we also use another jet substructure technique : N-subjettiness, which will be introduced in the following section. In summarization, these selections are applied to identify jets from hadronic W/Z decays:

- **Pruned jet mass m_{jet}** - Require the total pruned jet mass to satisfy $70 < m_{\text{jet}} < 100$ GeV.
- **N-subjettiness** - Require the 2-subjettiness/1-subjettiness (τ_2/τ_1) <0.5 for the **unpruned** CA8 jets.

A detailed performance study of this W-tagger has been made public in Refer-

ence [26].

3.3.1 N-subjettiness

N-subjettiness [28, 59, 60] exploits the fact that the pattern of the hadronic decay of a heavy object is reflected through the presence of distinctive energy lobes corresponding to the decay products, as opposed to QCD jets which present a more uniformly spread energy configuration (not aligned along the subjet axis). The inclusive jet shape variable N-subjettiness is defined, in its generalized version as derived in Reference [28], as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k}\, min((\Delta R_{1,k})^\beta, (\Delta R_{2,k})^\beta ... (\Delta R_{N,k})^\beta) \tag{3.1}$$

where the index k runs over the jet constituents and the distances $\Delta R_{n,k}$ are calculated with respect to the axis of the n^{th} subjet. The normalization factor d_0 is calculated as $d_0 = \sum_k p_{T,k} R_0^\beta$, setting R_0 to the jet radius of the original jet. In the analysis, the N-subjettiness is calculated from the unpruned jets with the parameter $\beta = 1$. In particular, the variable able to best discriminate between W/Z jets and QCD jets is the ratio of 2-subjettiness over 1-subjettiness, $\tau_{21} = \tau_2/\tau_1$, which turns out to be smaller for signal than for background as demonstrated in Figure 3.3.

We select "high purity" (HP) W/Z jets by requiring $\tau_{21} \leq 0.5$, while $0.5 < \tau_{21} < 0.75$ defines the "low purity" (LP) W/Z jets. The division of events with one W/Z-

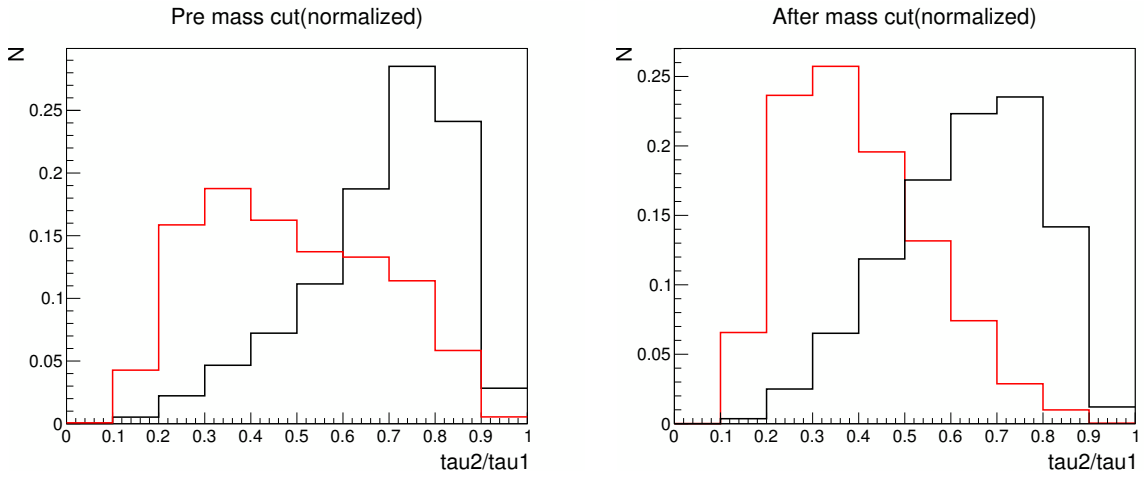

Figure 3.3: Comparison for τ_2/τ_1 distribution between signal (red) and background (black) before the jet mass cut (left) and after the jet mass cut applied (right). The signal MC used here is Herwig WW 1.5 TeV, and background is Herwig QCD.

tag follows the same delineation. The events with two W/Z-tagged jets are always required to have one HP W/Z tag, and are similarly divided into the HP and the LP categories depending on whether the other W/Z-tagged jet has passed the HP or the LP requirement, respectively. The HP category has been optimized to reach on average the best sensitivity for all models considered in this search. The LP category adds sensitivity in particular at high dijet masses where the W/Z-tagging efficiency drops along with the background rate.

3.3.2 Optimization study for the W-tagger

The cut values for the pruned jet mass and N-subjettiness were optimized based on the best expected limit. The final cut values are a compromise between best expected limits for WW and ZZ resonances in the range between 1 and 2 TeV, because we target

both of them with the same analysis.

Figure 3.4 shows the optimization of the N-subjettiness (τ_{21}) or massdrop (μ) cut value. The massdrop variable was used in the 2011 version of this analysis and has been replaced by the N-subjettiness. A N-subjettiness cut gives a 30% better limit than the massdrop cut. The $\tau_{21} < 0.5$ is the best cut value for equal performance in both WW and ZZ resonances. The expected limit changes by <5% changing the τ_{21} cut value by ± 0.05.

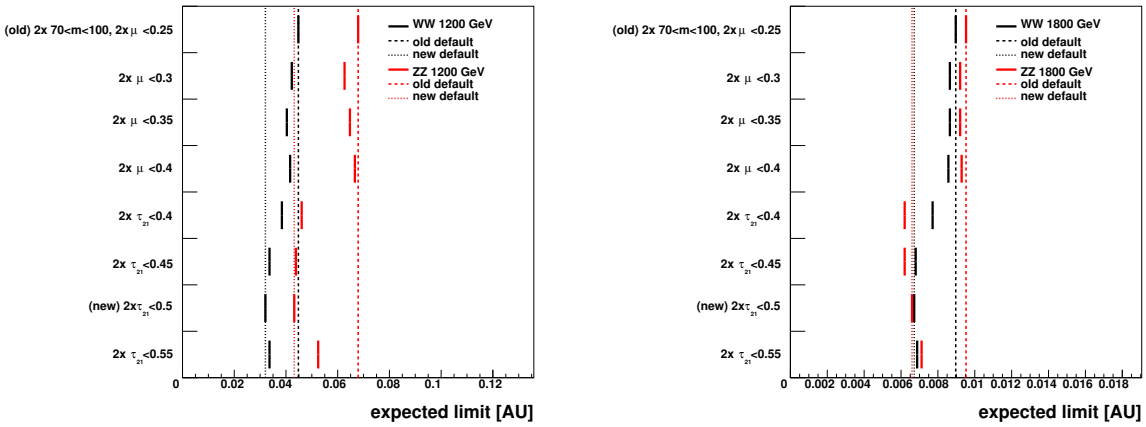

Figure 3.4: Optimizataion of the N-subjettiness (τ_{21}) or massdrop (μ) cut value for the best expected limit.Plot on the left hand is for HERWIGsamples of WW and ZZ at resonance mass 1.2 TeV , while 1.8 TeV for plot on the right hand.

Figure 3.5 shows the optimization of the pruned jet mass window cut. Neither widing nor narrowing the pruned jet mass window on either side can improve the expected limit for WW and ZZ at the same time. The jet mass window of $70 < m_{\text{jet}} < 100$ GeV provides best performance for WW and ZZ at the same time.

Figure 3.6 shows the dependency of the expected limit on the jet algorithm used

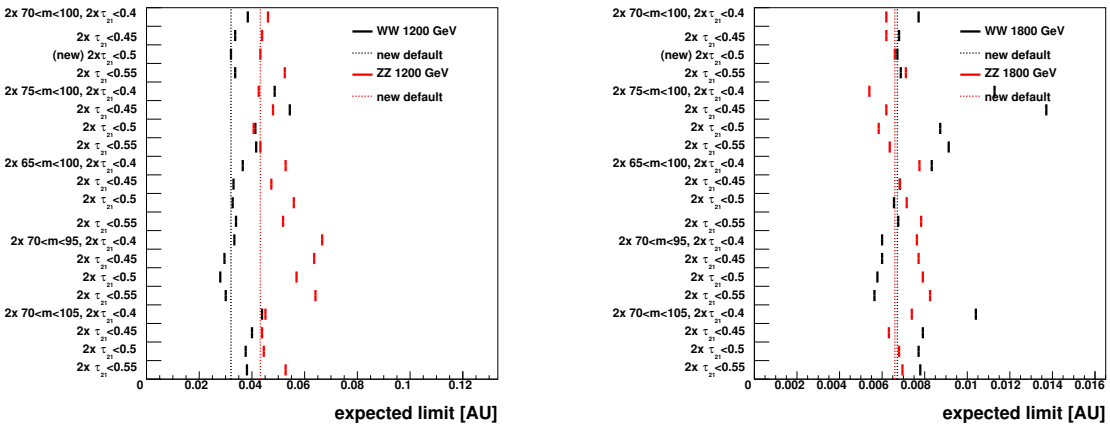

Figure 3.5: Optimizataion of the pruned jet mass window cut for the best expected limit.

for the resonance mass reconstruction. It is found that AK5, AK7 and CA8 show almost the same performance. This analysis switched since 2011 from AK5 to CA8 for consistency with other similar analyses.

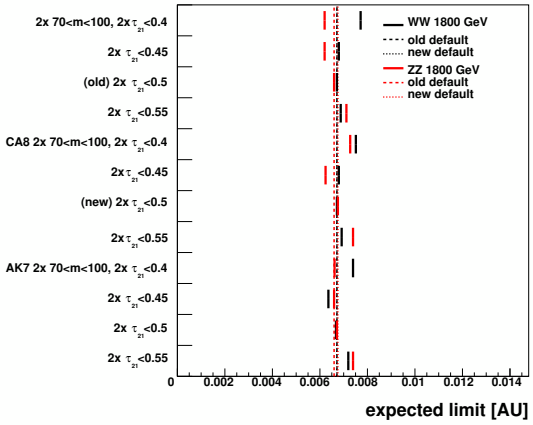

Figure 3.6: Comparison of expected limit for different jet algorithms.

3.4 Data and Monte Carlo samples

The data sample of proton-proton collisions at $\sqrt{s} = 8$ TeV was collected in 2012 and corresponds to an integrated luminosity of 19.7 fb^{-1}. The datasets is summarized in Table 3.1. The certification file used for these data is `Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt` . The dijet sample is dominated by light flavored and gluon jets, which we denote as the "QCD background". The estimation of QCD background is obtained from data by fitting an analytic parameterization of the dijet invariant mass distribution.

Dataset
/Jet/Run2012A-22Jan2013-v1/AOD
/JetHT/Run2012B-22Jan2013-v1/AOD
/JetHT/Run2012C-22Jan2013-v1/AOD
/JetHT/Run2012D-22Jan2013-v1/AOD

Table 3.1: Summary of 8 TeV collision data used in this analysis.

Signal events have been simulated using JHUGEN [61,62], PYTHIA 6.426 [63] and HERWIG++ 2.5.0 [64] event generators and processed through a simulation of the CMS detector, based on GEANT4 [65]. PYTHIA 6 is used with CTEQ6L1 [66] and HERWIG++ with MRST2001 [67] parton distribution functions. Tune Z2* (a modification of tune Z1 [68]) is used with PYTHIA 6, while the tune version 23 is used

with HERWIG++. The process $q^* \to$ W/Z + jet is generated using PYTHIA 6. RS graviton production is studied with $k/\overline{M}_{\rm Pl} = 0.1$, which determines a resonance width of about 1% of the resonance mass which is about a factor five smaller than the experimental resolution for dijets. While HERWIG++ contains a more detailed description of the angular distributions for $G_{\rm RS}$ than PYTHIA 6 for this process [69] and is therefore used to model the $G_{\rm RS}$ resonance shape, the PYTHIA 6 cross section is used to maintain consistency with reference models used in related analyses [36]. Bulk graviton production is studied with $k/\overline{M}_{\rm Pl} = 0.2$ and is generated with JHUGEN interfaced with PYTHIA 6 for the showering. Bulk graviton cross sections are calculated using CalcHEP. The process W$' \to$ WZ is generated using PYTHIA 6 with Standard Model $V - A$ couplings and without applying k-factors.

To validate our RS graviton Monte Carlo samples, we compare PYTHIA 6, HERWIG++ and JHUGEN (a generator including full angular correlations developed by the JHU group). Figure 3.7 shows the comparisions of invariant mass and $\Delta\eta$ of two Z bosons at generator level, in which HERWIG++ and PYTHIA 6 are compared with the JHU generator which describes the angular distributions exactly. PYTHIA 6 does not implement the angular correlations, and from Figure 3.7 one can indeed conclude that in its description of this effect it is inferior to HERWIG++.

All Monte Carlo events are fully simulated and reconstructed via the GEANT4-based CMS simulation and reconstruction software.

Tables 3.2, 3.3, 3.5 and 3.4 summarize the simulated signal samples used in this

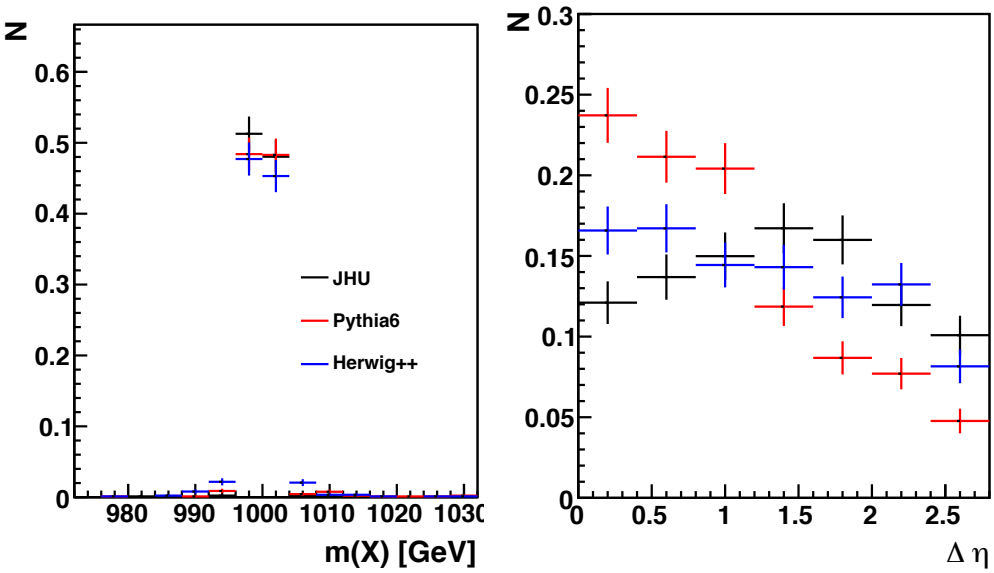

Figure 3.7: Invariant mass and $\Delta\eta$ of two Z bosons at generator level for comparing the PYTHIA 6 and the HERWIG++ models of a 1 TeV RS graviton resonance with $k/M_{PL} = 0.02$ with the JHU generator which includes all angular correlations.

analysis.

Table 3.2 describes a single-tagged process: $q* \to W/Z + \text{jet}$ with a large cross section. These samples are generated the MC using PYTHIA 6 with Tune Z2*. And the details of its configuration is in the Appendix A.1. The parameters `RTCM(43)`, `RTCM(44)`, `RTCM(45)` are set to 1 and the scale `RTCM(41)` is set to the resonance mass, `PMAS(343,1)`=`PMAS(344,1)`. Only decays into qW or qZ are allowed.

Table 3.3 shows a double-tagged process: $G_{RS} \to WW/ZZ$. This is produced using HERWIG++ with Tune23 and as a cross check also in PYTHIA 6 with Tune Z2*. In PYTHIA 6, the parameter `PARP(50)` corresponding to 5.4 $k/\bar{M}_{Pl}$ which impacts the width and cross section of the resonance. In HERWIG++, the cross section and

Process	Generator	Events	X-sec[pb]
qW(m=750 GeV)	PYTHIA 6	30000	1.133E+02
qW(m=1000 GeV)	PYTHIA 6	30000	2.647E+01
qW(m=1500 GeV)	PYTHIA 6	30000	2.540E+00
qW(m=2000 GeV)	PYTHIA 6	30000	3.510E-01
qW(m=3000 GeV)	PYTHIA 6	30000	1.008E-02
qZ(m=750 GeV)	PYTHIA 6	30000	4.071E+01
qZ(m=1000 GeV)	PYTHIA 6	30000	9.405E+00
qZ(m=1500 GeV)	PYTHIA 6	30000	8.937E-01
qZ(m=2000 GeV)	PYTHIA 6	30000	1.231E-01
qZ(m=3000 GeV)	PYTHIA 6	30000	3.465E-03

Table 3.2: Summary of the simulated Monte Carlo samples used in this analysis for the process q$*$ $\to$ Z/W + jet

width are given by the ratio of `RS/Model:Lambda_pi` and the resonance mass `/Herwig/Particles/Graviton:NominalMass`. The process $G_{RS} \to WW/ZZ$ is generated using Herwig++ with Tune23 and its cross section is taken from PYTHIA 6 with Tune Z2*. We study RS graviton production with $k/\bar{M}_{Pl} = 0.1$, defining a resonance width smaller than the experimental resolution for dijets. Table 3.4 describes another double-tagged process: $W' \to WZ$. This is produced using PYTHIA 6 with Tune Z2*. The decay of the W′ is restricted to WZ with `MDME(331,1)=1`.

Process	Generator	Events	PYTHIA 6 X-sec [pb]
WW(m=750 GeV)	HERWIG++/PYTHIA 6	30000	2.220E+00
WW(m=1000 GeV)	HERWIG++/PYTHIA 6	30000	4.254E-01
WW(m=1500 GeV)	HERWIG++/PYTHIA 6	30000	3.298E-02
WW(m=2000 GeV)	HERWIG++/PYTHIA 6	30000	4.083E-03
WW(m=2500 GeV)	HERWIG++/PYTHIA 6	30000	6.191E-03
WW(m=3000 GeV)	HERWIG++/PYTHIA 6	30000	1.010E-04
ZZ(m=750 GeV)	HERWIG++/PYTHIA 6	30000	1.120E+00
ZZ(m=1000 GeV)	HERWIG++/PYTHIA 6	30000	2.137E-01
ZZ(m=1500 GeV)	HERWIG++/PYTHIA 6	30000	1.662E-02
ZZ(m=2000 GeV)	HERWIG++/PYTHIA 6	30000	2.027E-03
ZZ(m=2500 GeV)	HERWIG++/PYTHIA 6	30000	3.077E-04
ZZ(m=3000 GeV)	HERWIG++/PYTHIA 6	30000	5.099E-05

Table 3.3: Summary of the simulated Monte Carlo samples used in this analysis for the process $G_{RS} \rightarrow WW, ZZ$.

Process	Generator	Events	X-sec[pb]
WZ(m=750 GeV)	PYTHIA 6	30000	5.391E-01
WZ(m=1000 GeV)	PYTHIA 6	30000	1.444E-01
WZ(m=1500 GeV)	PYTHIA 6	30000	1.804E-02
WZ(m=2000 GeV)	PYTHIA 6	30000	3.129E-03
WZ(m=2500 GeV)	PYTHIA 6	30000	6.781E-04
WZ(m=3000 GeV)	PYTHIA 6	30000	1.894E-04

Table 3.4: Summary of the simulated Monte Carlo samples used in this analysis for the process $W' \rightarrow WZ$.

Process	Generator	Events	X-sec[pb]
WW(m=1000 GeV)	JHU Z2*	50000	0.001774
WW(m=1500 GeV)	JHU Z2*	50000	9.207E-05
WW(m=2000 GeV)	JHU Z2*	50000	8.004E-06
WW(m=2500 GeV)	JHU Z2*	50000	8.851E-07
WW(m=3000 GeV)	JHU Z2*	50000	-
ZZ(m=1000 GeV)	JHU Z2*	50000	0.0009044
ZZ(m=1500 GeV)	JHU Z2*	50000	4.622E-05
ZZ(m=2000 GeV)	JHU Z2*	50000	4.029E-06
ZZ(m=2500 GeV)	JHU Z2*	50000	4.460E-07
ZZ(m=3000 GeV)	JHU Z2*	50000	-

Table 3.5: Summary of the simulated Monte Carlo samples used in this analysis for process $G_{Bulk} \to WW, ZZ$.

3.5 Trigger

Events are selected if one of the following triggers has fired: HLT_HT750, HLT_PFHT650, HLT_PFNoPUHT650, HLT_FatDiPFJetMass750_DR1p1_Deta1p5. All versions of each of these triggers is used. None of these triggers are prescaled during the 2012 data taking period. HLT_PFNoPUHT650 trigger is used for the data set after the RunC (including RunC), while HLT_PFHT650 trigger is only used for RunA and RunB data sets.

Figure 4.7, Figure 4.8, and Figure 4.9 show the trigger efficiencies of the OR of the highest threshold HLT_HT750, HLT_PFHT650, and the HLT_FatJetMass triggers w.r.t. an OR of the lower threshold HLT_HT550 trigger. From the plot, the trigger is 99% effiecient above 890 GeV for the untagged, single tagged and double tagged data.

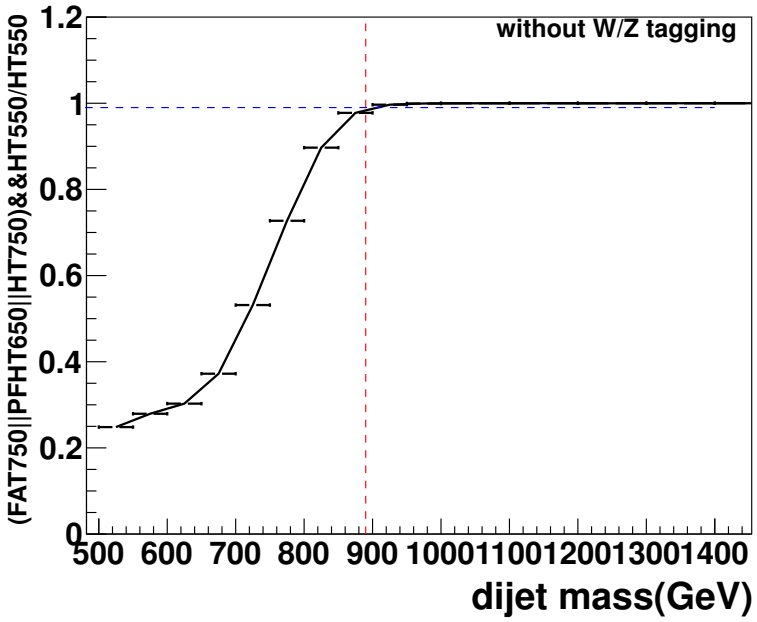

Figure 3.8: Trigger efficiency for untagged data of FAT_750‖HLT_PF(NoPU)HT650‖HLT_HT750 measured using data collected by lower threshold HLT_HT550 trigger. The dash red line is positioned at m_{jj} equal 890 GeV , the blue line is at efficiency at 99%.

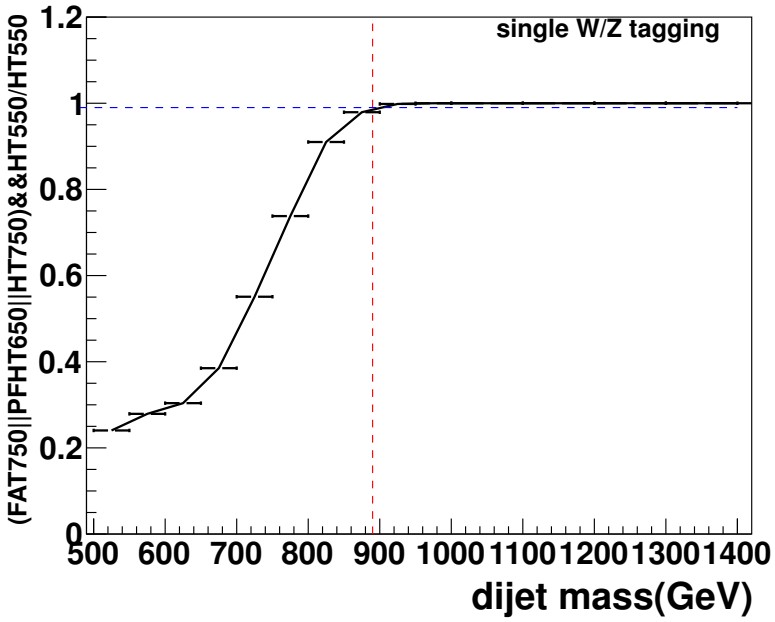

Figure 3.9: Trigger efficiency for single tagged data of FAT_750||HLT_PF(NoPU)HT650||HLT_HT750 measured using data collected by lower threshold HLT_HT550 trigger. The dash red line is positioned at m_{jj} equal 890 GeV , the blue line is at efficiency at 99%.

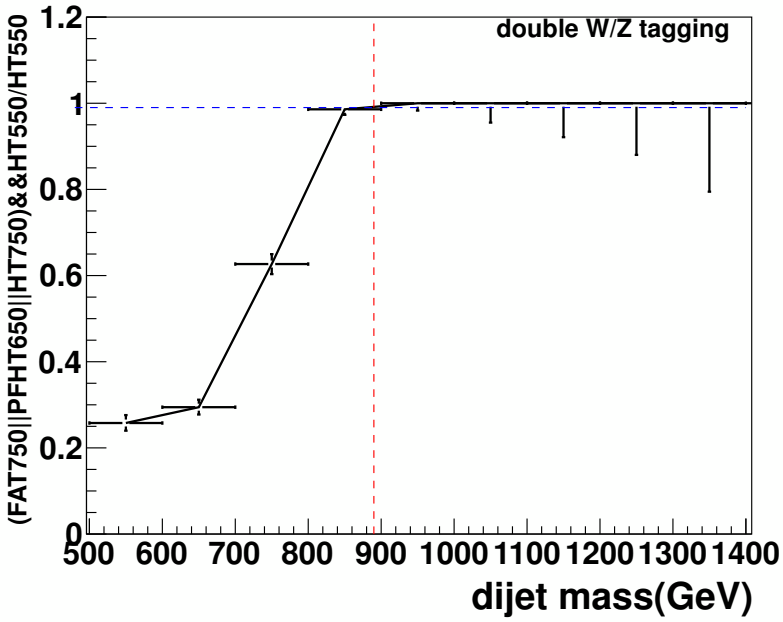

Figure 3.10: Trigger efficiency for double tagged data of FAT_750||HLT_PF(NoPU)HT650||HLT_HT750 measured using data collected by lower threshold HLT_HT550 trigger. The dash red line is positioned at m_{jj} equal 890 GeV , the blue line is at efficiency at 99%.

3.6 Data and MC comparisons

In this section, we compare some kinematic features of the jets between QCD MC and data, which are shown in Figure 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, 3.17, 3.18, 3.19, 3.20, and 3.21. Predictions from PYTHIA 6 with Tune Z2* and HERWIG++ with Tune 23 are shown. The comparison is shown in the exclusive dijet category (LP and HP), single and double tagged events. The distributions are shown after the event selection (in particular $|\eta| < 2.5$, $|\Delta\eta| < 1.3$). The number of data events in each mass bin is shown in Table 3.6. The MC is normalized to the number of data events in each category and the shapes are compared.

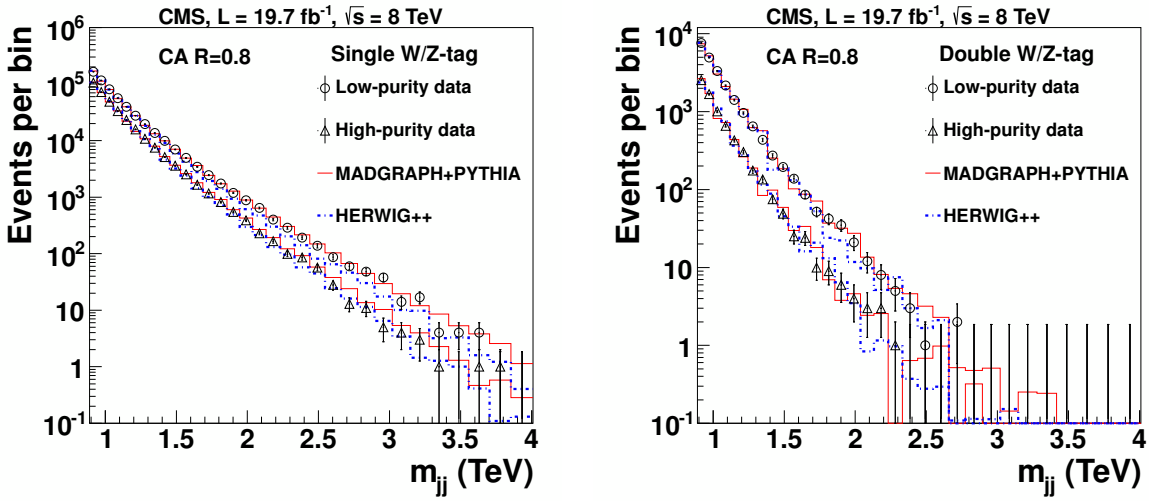

Figure 3.11: The m_{jj} distributions for (left) singly and (right) doubly tagged events in data, and QCD multijet (MADGRAPH/PYTHIA and HERWIG++) simulations, normalized to data.

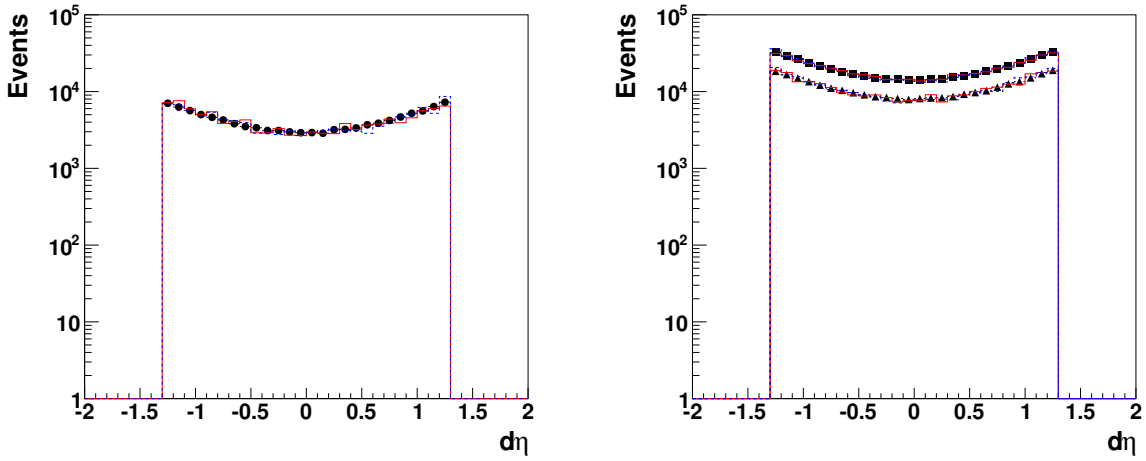

Figure 3.12: Comparisons between data and Monte Carlo for $\Delta\eta$ of the two leading jets of failed tagging events (left) and LP&HP purity (right) 1-tagged events. The MC is normalized to the number of data events in each category.

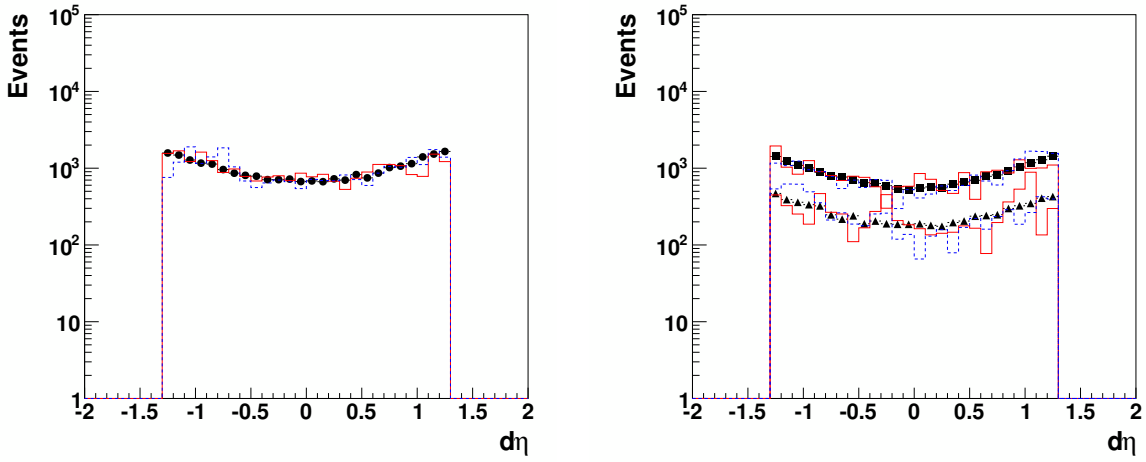

Figure 3.13: Comparisons between data and Monte Carlo for $\Delta\eta$ of the two leading jets of failed tagging events (left) and LP&HP purity (right) 2-tagged events. The MC is normalized to the number of data events in each category.

	1-tag events		2-tag events	
lower mass bin boundary	LP	HP	LP	HP
2037	643	230	12	3
2132	402	167	8	3
2231	287	99	5	1
2332	193	86	3	
2438	138	57	1	
2546	87	28	0	
2659	60	13	2	
2775	48	11		
2895	38	5		
3019	14	4		
3147	17	3		
3279	4	1		
3416	4	0		
3558	4	1		
3704		1		
3854				
4010				

Table 3.6: Number of events in each mass bin exclusive, with 1 W/Z-tag and 2 W/Z-tags required in LP and HP categories for events with resonance masses > 2 TeV.

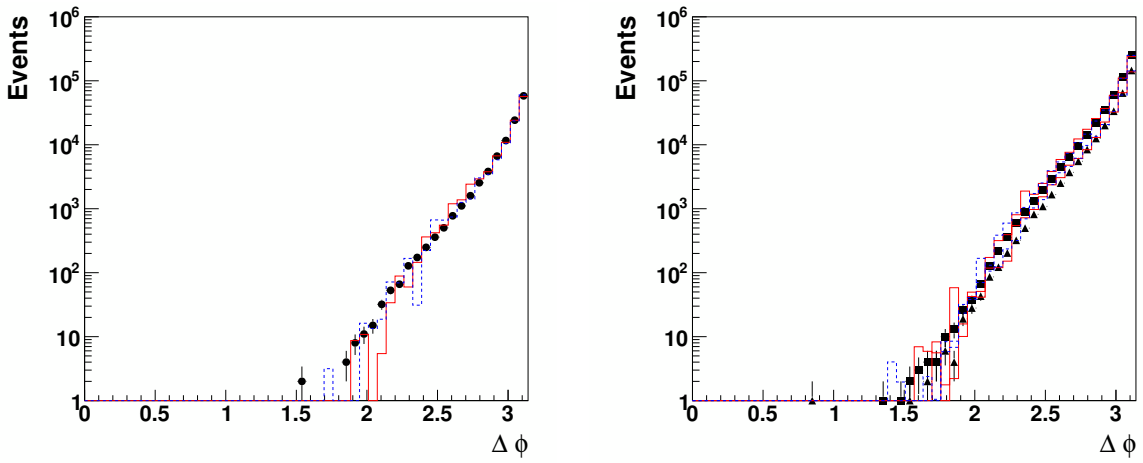

Figure 3.14: Comparisons between data and Monte Carlo for $\Delta\phi$ of the two leading jets of failed tagging events (left) and LP&HP purity (right) 1-tagged events. The MC is normalized to the number of data events in each category.

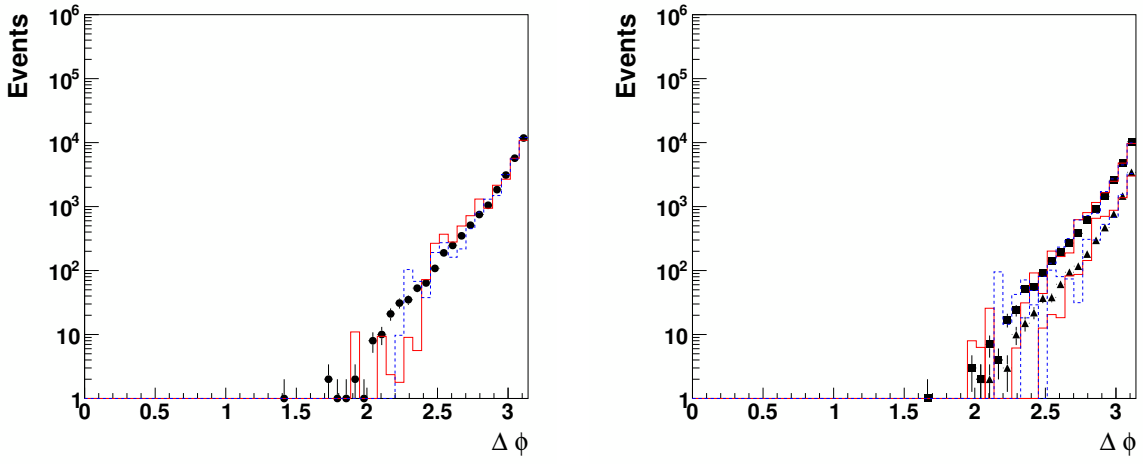

Figure 3.15: Comparisons between data and Monte Carlo for $\Delta\phi$ of the two leading jets of failed tagging events (left) and LP&HP purity (right) 2-tagged events. The MC is normalized to the number of data events in each category.

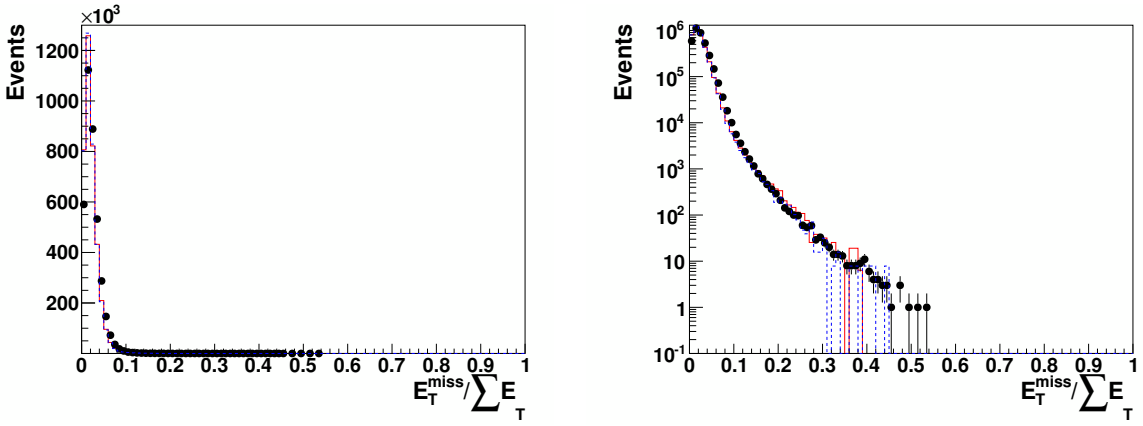

Figure 3.16: Comparisons between data and Monte Carlo for $E_T^{miss}/\sum E_T$. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

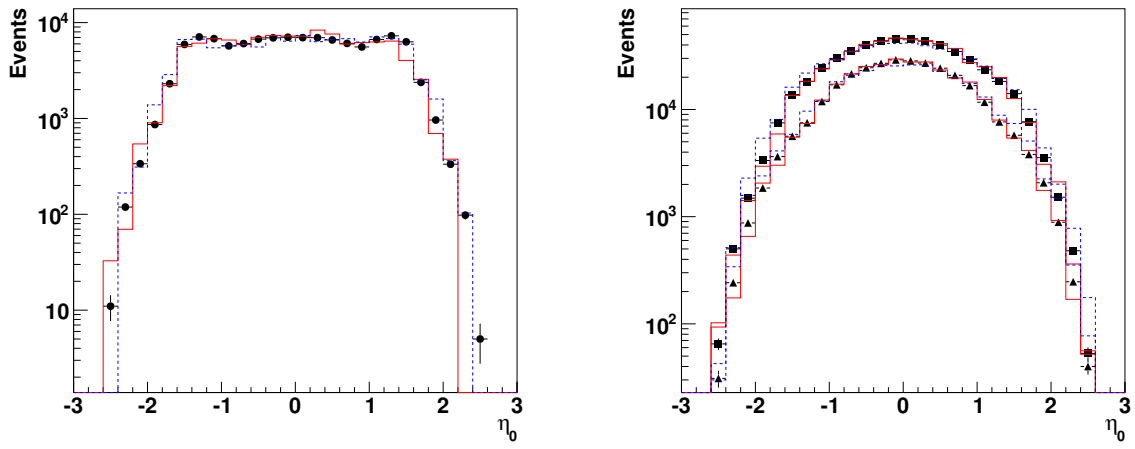

Figure 3.17: Comparisons between data and Monte Carlo for η of the leading jet of failed tagging events (left) and LP&HP purity (right) 1-tagged events. The MC is normalized to the number of data events in each category.

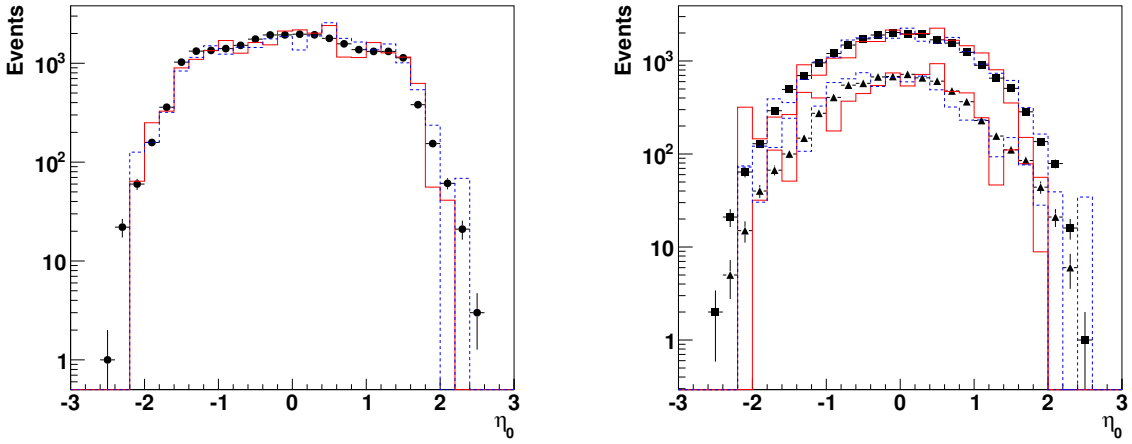

Figure 3.18: Comparisons between data and Monte Carlo for η of the leading jet of failed tagging events (left) and LP&HP purity (right) 2-tagged events. The MC is normalized to the number of data events in each category.

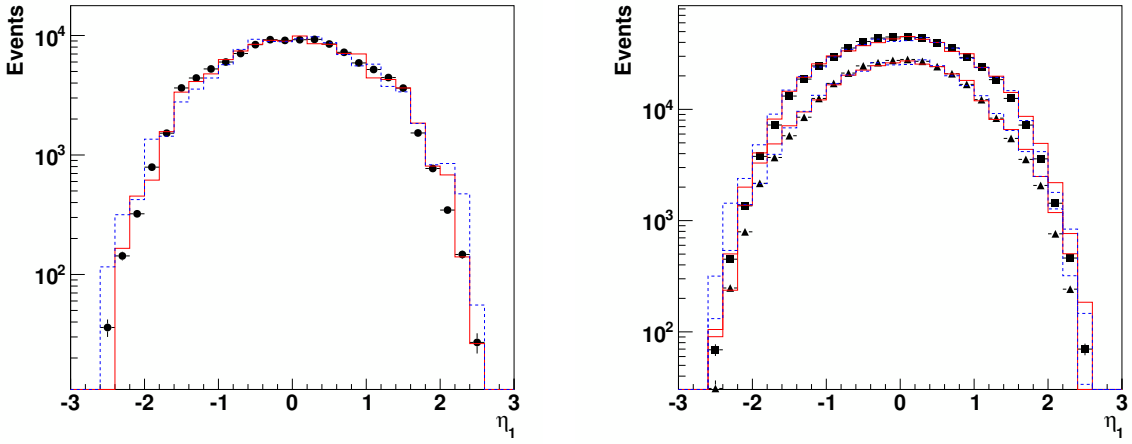

Figure 3.19: Comparisons between data and Monte Carlo for η of the second leading jet of failed tagging events (left) and LP&HP purity (right) 1-tagged events. The MC is normalized to the number of data events in each category.

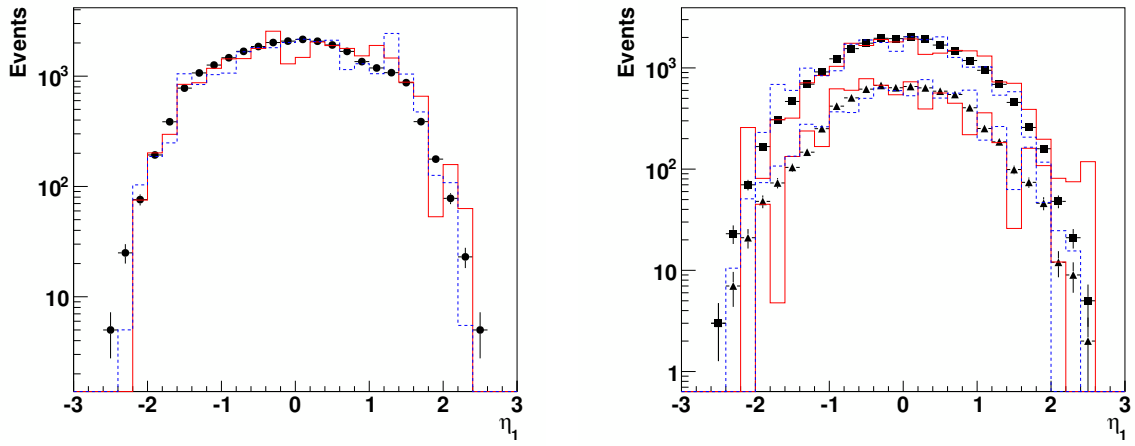

Figure 3.20: Comparisons between data and Monte Carlo for η of the second leading jet of failed tagging events (left) and LP&HP purity (right) 2-tagged events. The MC is normalized to the number of data events in each category.

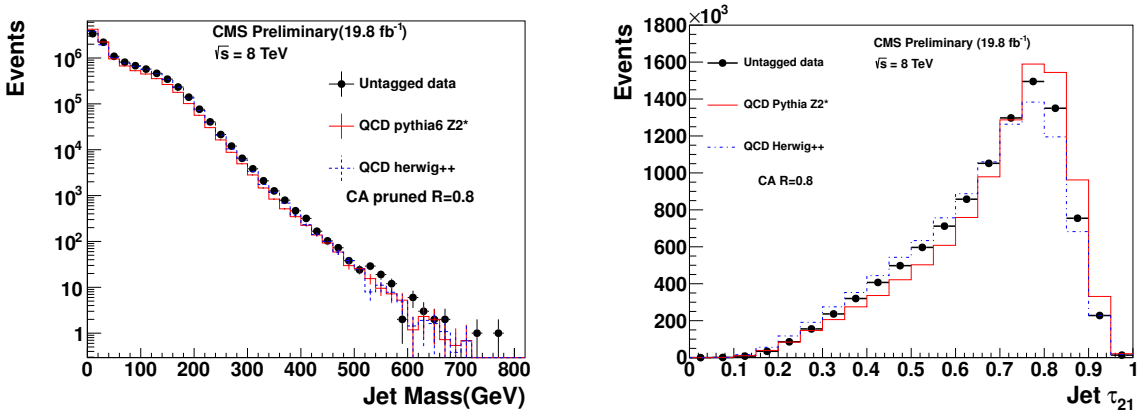

Figure 3.21: Comparisons between data and Monte Carlo for mass(left) and τ_{21}(right) of the leading two jets. The MC is normalized to the number of data events in each category.

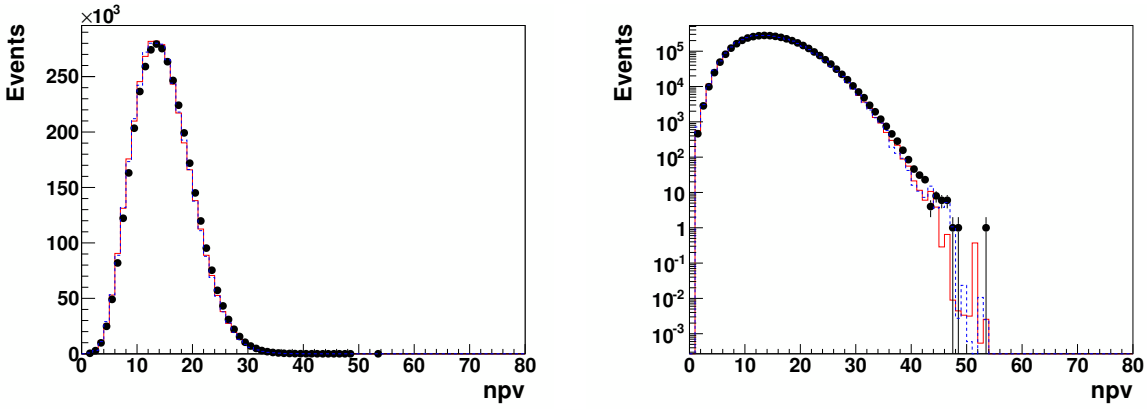

Figure 3.22: Comparisons between data and Monte Carlo for number of primary vertices to show the effect on Monte Carlo after pile up reweighting. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

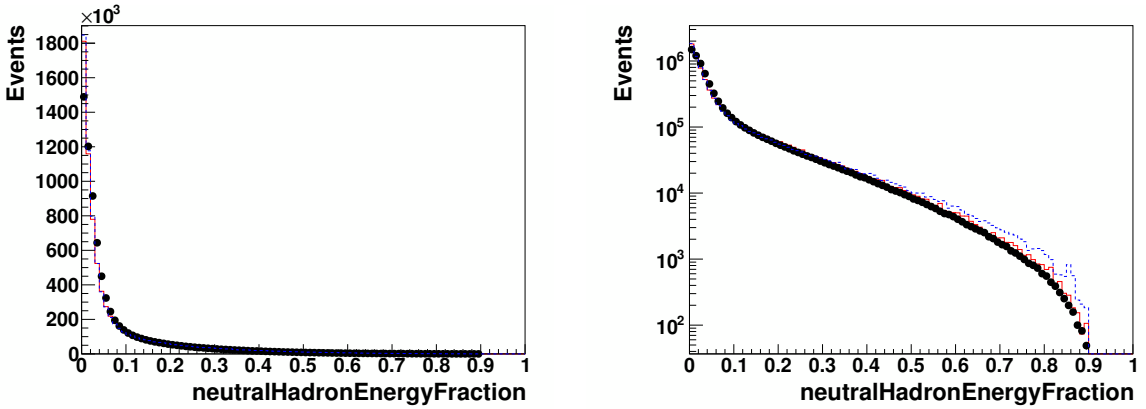

Figure 3.23: Comparisons between data and Monte Carlo for neutral hadron energy fraction. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

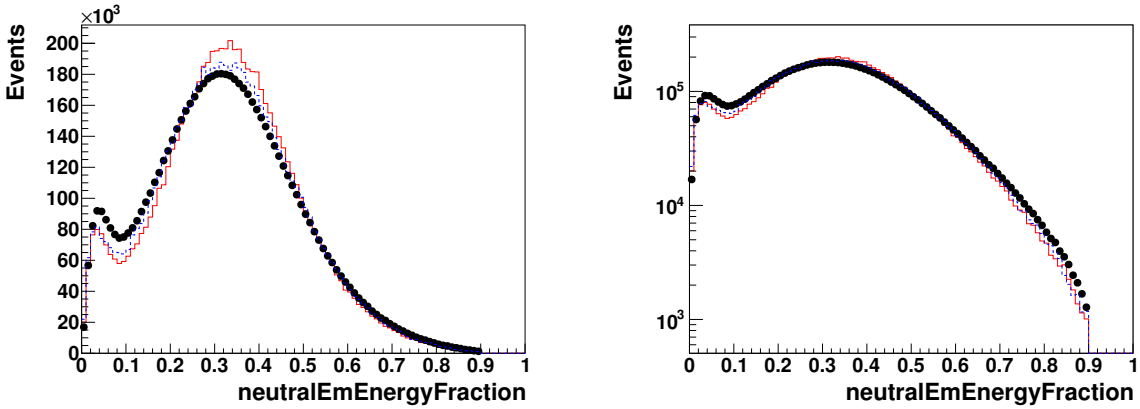

Figure 3.24: Comparisons between data and Monte Carlo for neutral eletromagnetic energy fraction. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

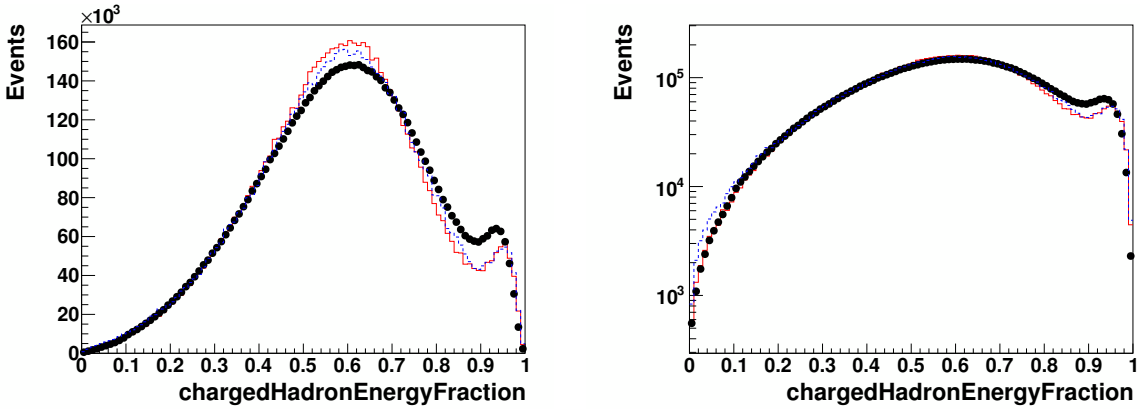

Figure 3.25: Comparisons between data and Monte Carlo for charged hadron energy fraction. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

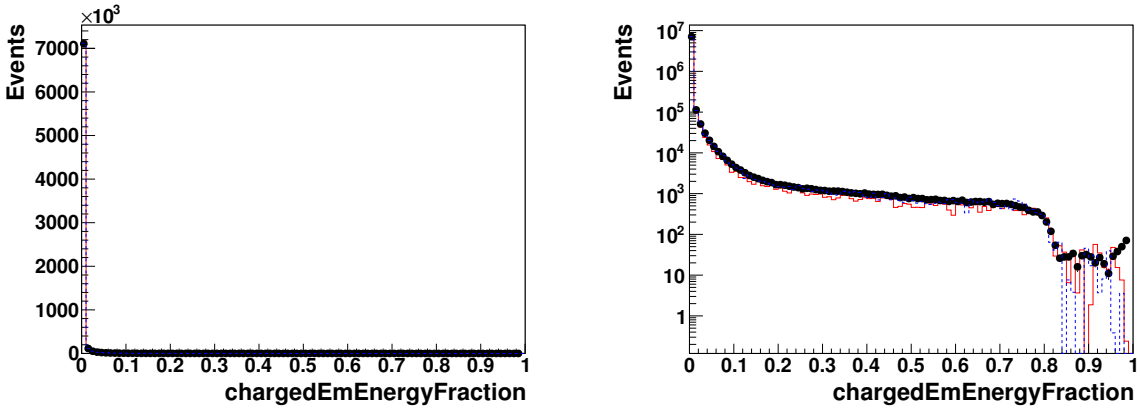

Figure 3.26: Comparisons between data and Monte Carlo for charged eletromagnetic energy fraction. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

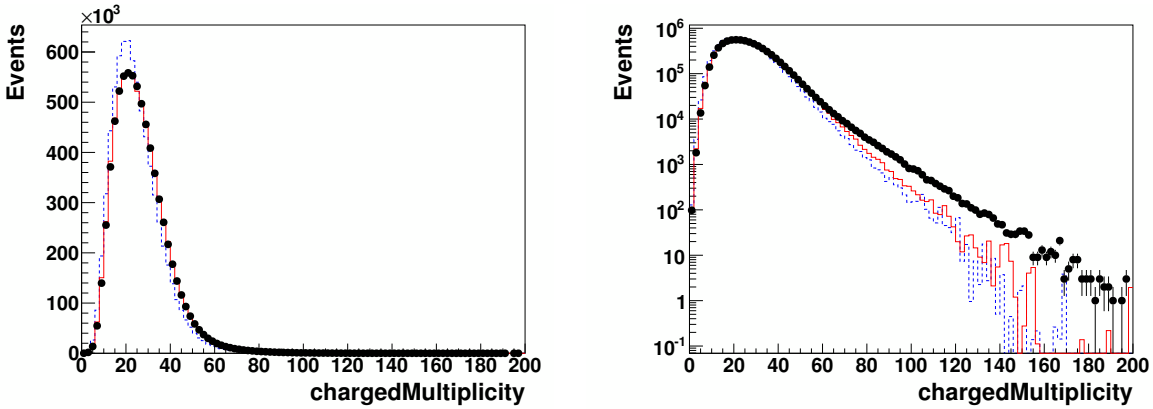

Figure 3.27: Comparisons between data and Monte Carlo for charged multiplicity. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

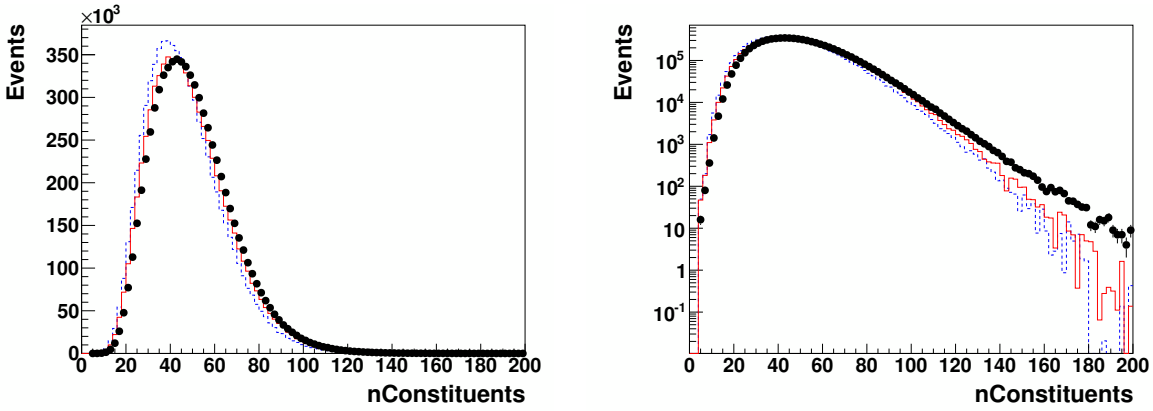

Figure 3.28: Comparisons between data and Monte Carlo for number of constituents. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

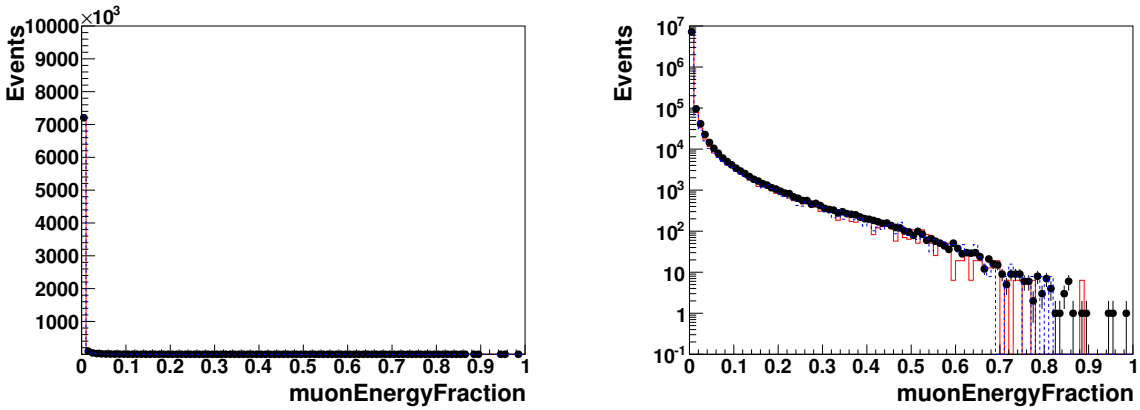

Figure 3.29: Comparisons between data and Monte Carlo for the muon energy fraction of the leading two jets. The MC is normalized to the number of data events. Plot on the right is the log scale plot. (The plot includes only a subset of the full data sample.)

We find that the QCD MC agrees with data, although not perfect. For the dijet kinematics and also the jet substructure variables, we observe about the same agreement of PYTHIA 6 and HERWIG++. For this analysis, we chose to model the background shape from the data itself (as described below) and depend on QCD MC only to provide us guidance and a cross check.

Figure 3.14, Figure 3.15 and Figure 3.16 are particulariy useful to identify jets from calorimeter noise which would show up at low values of $\Delta\phi$ and high values of $E_T^{miss}/\sum E_T$. No enhancement in this region is observed which gives confidence that the applied noise filters and jet ID cuts leave no noise contamination within the two leading jets.

Figure 3.22 shows the number of primary vertices distribution after pile up reweighting on the MC. Figures 3.23, 3.24, 3.25, 3.26, 3.27 and 3.28 show the jet ID variable

distribution after the event selection, and Figure 3.29 shows the muon energy faction of the leading two jets.

3.7 The signal: dijet resonance

We search for dijet resonances corresponding to several models. Using the W/Z-tagging algorithm, we examine both single W/Z-tag and double W/Z-tag events.

The pruned jet mass and jet τ_{21} distributions in signal MC, data and background MC are shown in Figure 3.30. Fully merged jets from hadronic W and Z decays peak around 80∼90 GeV while QCD jets and not fully merged W and Z jets peak around 20 GeV. The distribution of τ_{21} for W/Z $\to$ qq′ signal, peaks around 0.4 and is almost fully contained within $\tau_{21} < 0.75$, where we place our cut; in contrast, QCD background peaks $\sim 0.7 - 0.8$. The discriminating power of the pruned jet mass and τ_{21} is evident.

The modelling of the signal efficiency is cross-checked through a W-tagging efficiency estimated using merged W $\to$ qq′ decays in $t\bar{t}$ events [26]. The efficiency is obtained using ℓ + jets events with two b-tagged jets, one of which has $p_T > 200$ GeV. Such events are dominated by $t\bar{t}$ production. The data are compared to simulated $t\bar{t}$ events, generated with MADGRAPH, interfaced to PYTHIA for parton showering, and provide scale factors of 0.86 ± 0.07 and 1.39 ± 0.75, respectively, for HP and LP events. These values are derived following the method described in Ref [26] for the selections applied in this analysis, and are used to match the simulated samples to data. The uncertainties in the scale factors contribute to the systematic uncertainty in the selection efficiency for signal.

For both the pruned jet mass and τ_{21}, differences are observed between the HER-

WIG++ (G_{RS}) and PYTHIA 6 (G_{Bulk}, q^*, W′) distributions, which arise from differences in the polarization of the W/Z boson and the showering and hadronization models used by these generators. The differences, due to showering and hadronization, are taken into account in estimating the systematic uncertainties on the tagging efficiencies, as discussed below.

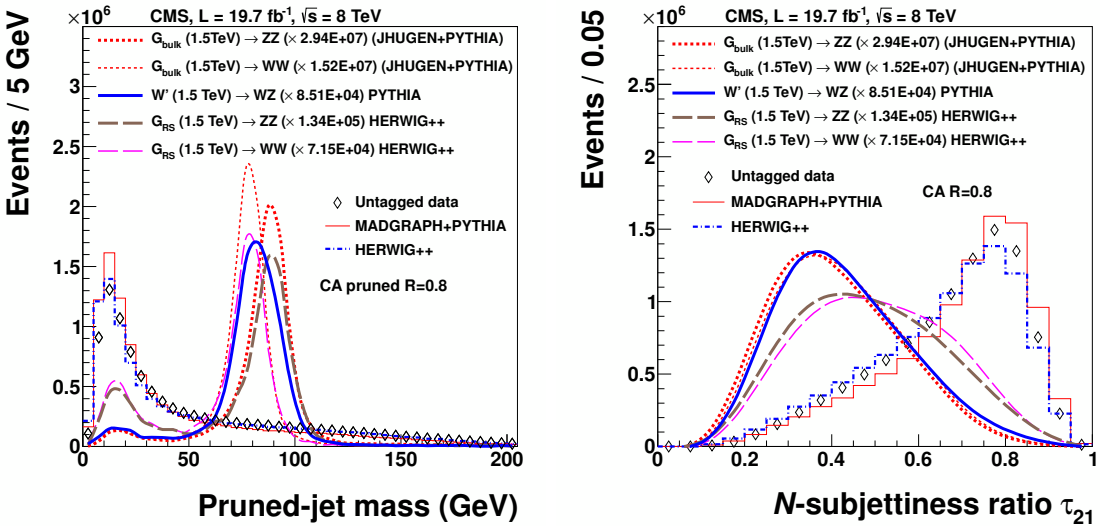

Figure 3.30: Distribution for (left) pruned-jet mass m_j and (right) jet N-subjettiness ratio τ_{21} in data, and in simulations of signal and background events. All simulated distributions are scaled to match the number of events in data. MADGRAPH/PYTHIAand HERWIG++ refer to QCD multijet event simulations.

The full event selection efficiency is estimated using simulated signal samples. Less than 1% of the ZZ or WW events which pass the full selection are from ZZ $\to llqq$ or WW $\to l\nu qq$ decays, where l can be a muon or electron. While 3% of the selected ZZ events are from ZZ $\to \tau\tau qq$ decays, less than 1% of the selected WW events are from WW $\to \tau\nu qq$ decays. To within 10% accuracy, the full selection efficiency can therefore be approximated by the product of the W/Z-tagging efficiency and an

approximate acceptance. This acceptance is shown in Figure 3.31 and takes into account the angular acceptance ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$), the branching faction into quark final states $\mathcal{B}(\mathrm{W/Z} \to \mathrm{qq}\prime)$, and a matching within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.5$ between the generated W/Z bosons and the reconstructed jets.

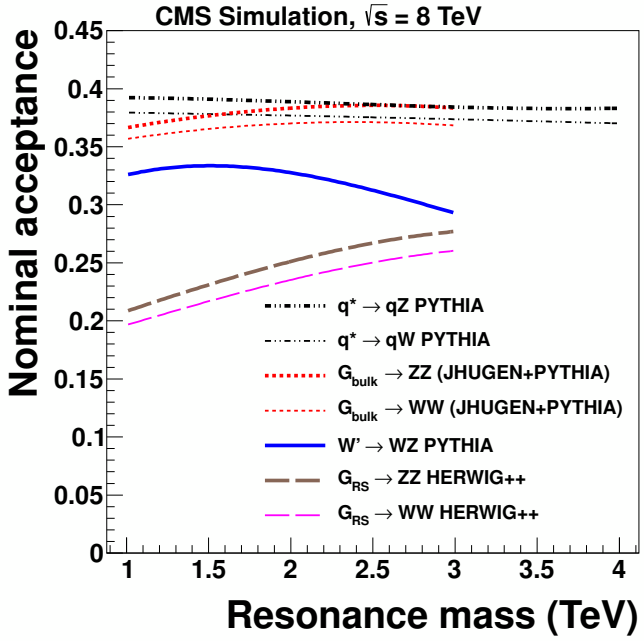

Figure 3.31: The fraction of simulated signal events expected for vector bosons decaying into two quarks, reconstructed as two jets, that pass the geometrical acceptance criteria ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$), shown as a function of the dijet invariant mass.

The W/Z-tagging efficiency and also the tag rate in data are shown in Figure 3.32. Since data is dominated by background events, the tag rate in data could be viewed as mistag rate.

The signal shapes of the HP category for all the processes considered in this analysis are shown in Figure 3.33. For the qW and qZ final states, the signal shapes

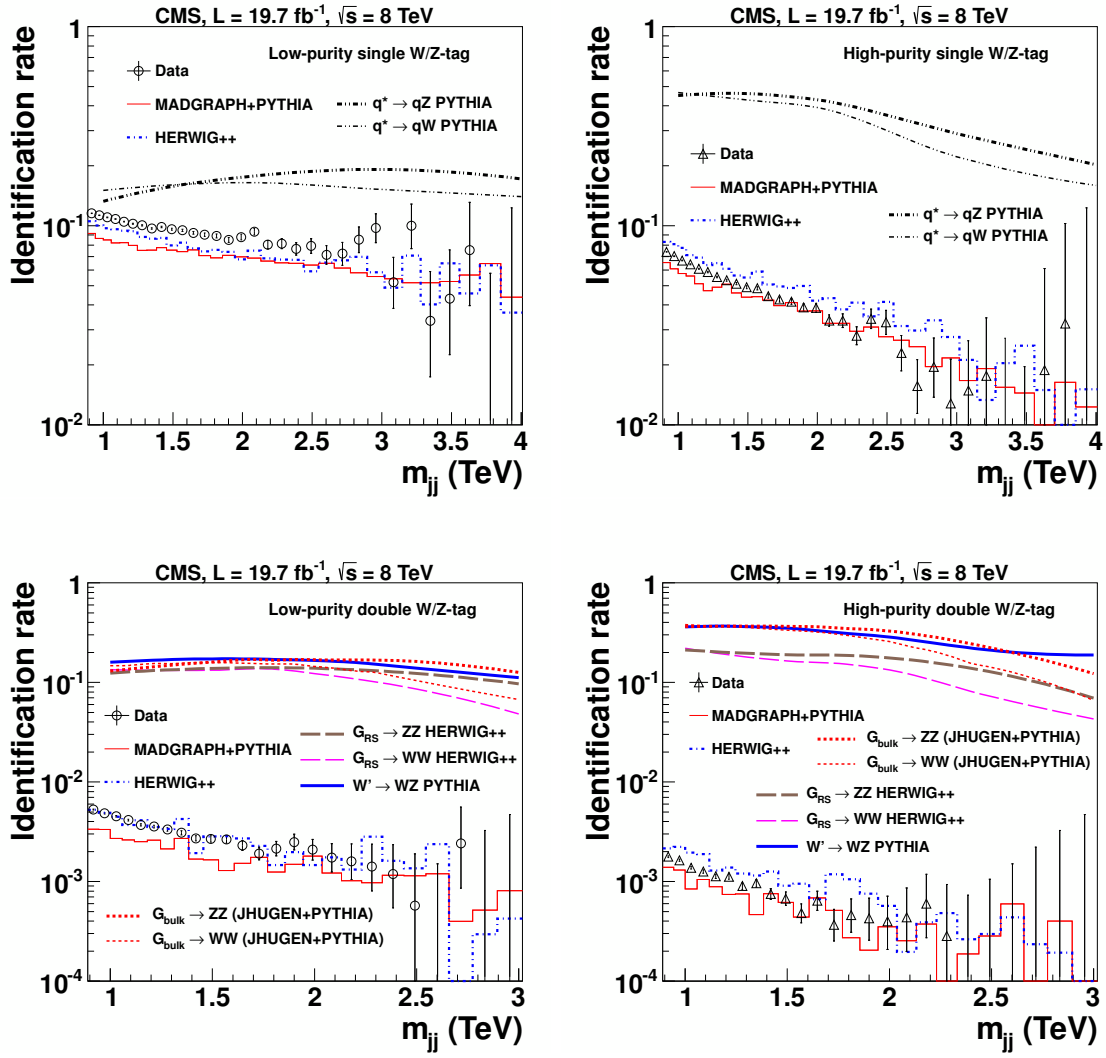

Figure 3.32: Identification rate for W and Z boson selections as a function of m_{jj} for quark and gluon jets in data and in simulation of background events, and for jets from W and Z bosons in simulation of signal events, with (upper left) one LP or (upper right) HP W/Z-tag, and the fraction of (lower left) doubly-tagged events in the LP and (lower right) HP category. The identification rate is computed for W/Z $\to q\bar{q}' \to$ jets events, where the jets have $|\eta| < 2.5$ and $|\Delta\eta| < 1.3$. MADGRAPH/PYTHIAand HERWIG++ refer to QCD multijet event simulations.

with a single W/Z-tag required are shown, while for the other signals two W/Z-tags are required.

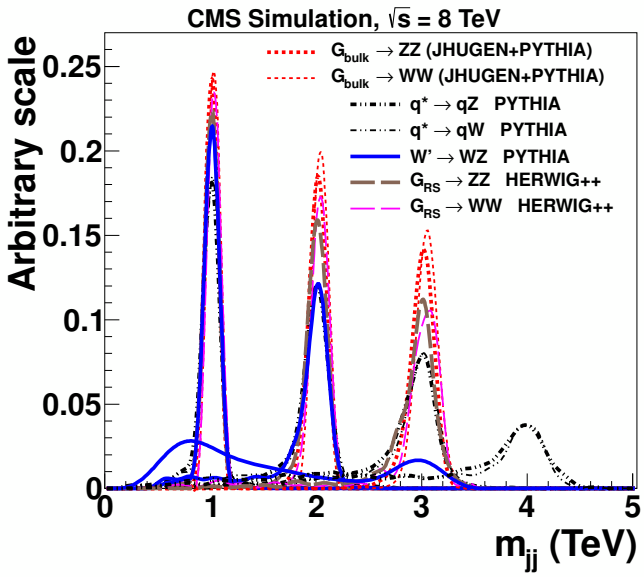

Figure 3.33: The normalized HP signal resonance distribution for $G_{RS} \to WW$, $G_{RS} \to ZZ$, $W' \to WZ$, $q* \to qW$, and $q* \to qZ$ resonances of dijet invariant mass 1.0 TeV , 1.5 TeV , 2.0 TeV , 3.0 TeV , 4.0 TeV .

3.8 Systematic uncertainties

The sources of systematic uncertainties are summarized as follows:

- Background-related systematic uncertainties: background shape parametrization.
- Signal-related systematics uncertainties: the determination of W/Z-tagging efficiency, jet energy scale(JES), jet energy resolution(JER), luminosity, PDF, and pile up.

3.8.1 Background shape parametrization

We model the shape of the QCD background in the dijet spectrum using a simple parametrization which has been successfully deployed in previous searches in the dijet mass spectrum [70]. Note in the limit setting, we employ a background plus signal fit. Here we show the background only fit to prove that data is dominated by background. The background model is given in Equation (3.2):

$$\frac{\mathrm{d}N}{\mathrm{d}m} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}}. \tag{3.2}$$

where m denotes the dijet mass and $\sqrt{s}$ is the center of collision energy for pp process. P_0 acts as a normalization parameter for the probability density function, and P_1, P_2 describe its shape. It has been checked by a Fisher F-test that no additional

parameter is needed to describe the distributions.

Figure 3.34 show the dijet mass spectra for single and double W/Z-tagged data, fitted to Equation (3.2) and the bottom panes show corresponding pull distributions, demonstrating the agreement between the background-only probability density function and the data.

No sizable deviation from the background-only hypothesis is seen, exclusion limits are set on the product of cross section, acceptance, and branching fraction for the five considered final states: qW, qZ, WW, WZ, and ZZ.

3.8.2 W/Z-tagging efficiency

The uncertainty in the efficiency for singly W/Z-tagged events is estimated using the ℓ+jets control sample from $t\bar{t}$ events described above. Uncertainties of 7.5% and 54% in the respective scale factors for HP and LP tagging include contributions from control-sample statistical uncertainties, and the uncertainties in the JES and JER for pruned jets. Since the scale factors are estimated only in the kinematic regime of the $t\bar{t}$ sample, where the W decay products merge and the b quarks are reconstructed as separate jets, we use the simulation just to extrapolate to larger W/Z-jet p_T. The efficiency is therefore estimated as a function of p_T for two showering and hadronization models, using G_{Bulk} samples generated with the JHUGEN event generator interfaced to PYTHIA and HERWIG++. The differences are respectively within 4% and 12% for HP and LP tagged jets, significantly smaller than the statistical

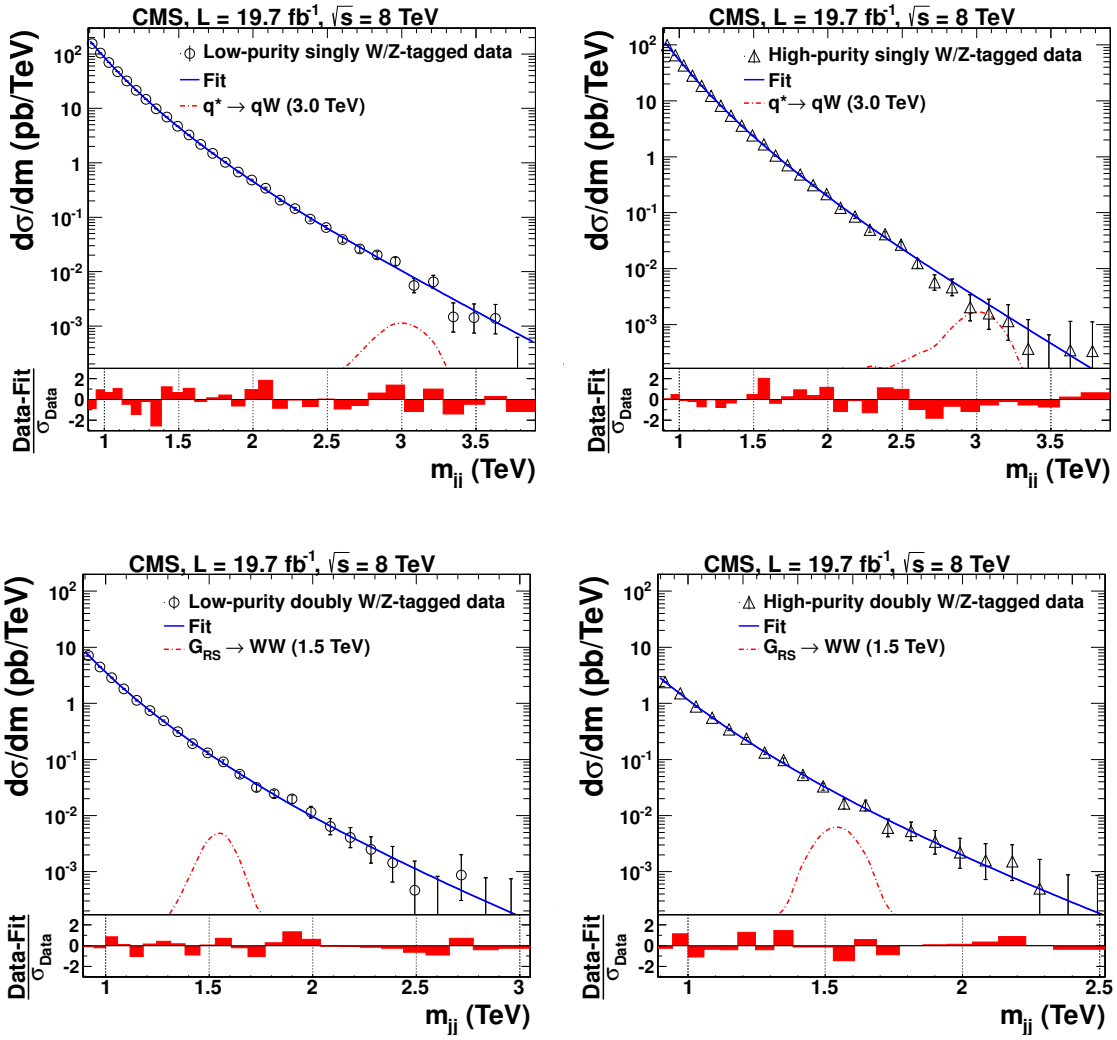

Figure 3.34: Distribution in m_{jj}, respectively, for (upper left) singly-tagged LP events and (upper right) HP events, and for (lower left) doubly-tagged LP events and (lower right) HP events. The solid curves represent the results of fitting Eq. (4.1) to the data. The distribution for $q^* \to qW$ and $G_{RS} \to WW$ contributions, scaled to their corresponding cross sections, are given by the dash-dotted curves. The corresponding pull distributions ($\frac{\text{Data}-\text{Fit}}{\sigma_{\text{Data}}}$, where σ_{Data} represents the statistical uncertainty in the data in a bin in m_{jj}) are shown below each m_{jj} plot.

uncertainties in the scale factors. Other systematic uncertainties in tagging efficiency are even smaller. Because of the rejection of charged particles not originating from

the primary vertex, and the application of pruning, the dependence of the W/Z-tagging efficiency on pileup is weak, and the uncertainty in the modelling of the pileup distribution is <1.5%. These systematic contributions refer to a singly W/Z-tagged jet, and are applied to each of the two leading jets in doubly W/Z-tagged events.

The JES has an uncertainty of 1–2% [54, 71], and its p_T and η dependence is propagated to the reconstructed value of m_{jj}, yielding an uncertainty of 1%, regardless of the resonance mass. The impact of this uncertainty on the calculated limits is estimated by changing the dijet mass in the analysis within its uncertainty. The JER is known to a precision of 10%, and its non-Gaussian features observed in data are well described by the CMS simulation [54]. The effect of the JER uncertainty in the limits is also estimated by changing the reconstructed resonance width within its uncertainty. The integrated luminosity has an uncertainty of 2.6% [72], which is also taken into account in the analysis. The uncertainty related to the PDF used to model the signal acceptance is estimated from the eigenvectors of the CTEQ66 [66] and MRST2006 [73] sets of PDF. The envelope of the upward and downward variations of the estimated acceptance for the two sets is assigned as uncertainty and found to be 5% – 15% in the resonance mass range of interest. A summary of all systematic uncertainties is given in Table 3.7.

Table 3.7: Summary of systematic uncertainties. The labels HP and LP refer to high-purity and low-purity event categories, respectively.

Source	Relevant quantity	LP (%)	HP (%)
Jet energy scale	Resonance shape	1	1
Jet energy resolution	Resonance shape	10	10
W-tagging	Efficiency (per jet)	54	7.5
Tagging p_T-dependence	Efficiency (per jet)	<4	<12
Pileup	Efficiency (per jet)	<1.5	<1.5
Integrated luminosity	Yield (per event)	2.6	2.6
PDF	Yield (per event)	5–15	5–15

3.9 Limit setting procedure

We search for a peak on top of the falling background spectrum by means of a maximum likelihood fit to the data. The likelihood $\mathcal{L}$, computed using events binned as a function of m_{jj}, is written as

$$\mathcal{L} = \prod_i \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}, \tag{3.3}$$

where $\lambda_i = \mu N_i(S) + N_i(B)$, μ is a scale factor for the signal, $N_i(S)$ is the number expected from the signal, and $N_i(B)$ is the number expected from multijet background. The parameter n_i quantifies the number of events in the i^{th} m_{jj} mass bin. The background $N_i(B)$ is described by the functional form of Equation (3.2). While maximizing the likelihood as a function of the resonance mass, μ as well as the parameters of the background function are left floating.

We quantify the consistency of the data with the null hypothesis as a function of resonance mass for the benchmark models through the local p-value. The largest local significance in the singly W/Z-tagged sample is observed for the hypothesis of a q$* \to$ qW resonance of mass 1.5 TeV, and is equivalent to an excess of 1.8 standard deviations. The largest local significance in the doubly tagged event sample corresponds to an excess of 1.3 standard deviations for a $G_{\mathrm{RS}} \to$ WW resonance of mass 1.9 TeV. Using the $\mathrm{G_{Bulk}} \to$ WW/ZZ model, where the LP and HP categories contribute in different proportions compared to the case for the $G_{\mathrm{RS}} \to$ WW model,

yields no excess larger than one standard deviation.

Using pseudo-experiments, we estimated the probability of observing a local statistical fluctuation of at least two standard deviations in any mass bin. This probability corresponds to an equivalent global significance of one standard deviation. The m_{jj} distributions are used to set upper limits on the product of the production cross sections and decay branching fractions for the benchmark models.

3.10 Results

The asymptotic approximation [74] of the LHC CL_s method [75, 76] is used to set upper limits on the cross sections for resonance production. The dominant sources of systematic uncertainties are treated as nuisance parameters associated with log-normal priors in those variables, following the methodology described in Reference [77]. For a given value of the signal cross section, the nuisance parameters are fixed to the values that maximize the likelihood, a method referred to as profiling. The dependence of the likelihood on parameters used to describe the background in Equation (3.2) is removed in the same manner, and no additional systematic uncertainty is therefore assigned to the parameterization of the background.

The HP and LP event categories are combined into a common likelihood, with the two uncertainties in the W/Z-tagging efficiencies considered to be anticorrelated between HP and LP tagging because of the exclusive selection on τ_{21}, while the remaining systematic uncertainties in signal are taken as fully correlated. The variables describing the background uncertainties are treated as uncorrelated between the two categories. The LP category contributes to the sensitivity of the analysis, especially at large values of m_{jj}. The combined expected limits on the $G_{\mathrm{RS}} \to \mathrm{WW}$ production cross sections are, respectively, a factor of 1.1 and 1.6 smaller at $m_{\mathrm{jj}} = 1.0$ TeV and 2.9 TeV than the limit obtained from the HP category alone.

Figures 3.35 and 3.36 show the observed and background-only expected upper limits on the production cross sections for singly and doubly W/Z-tagged events,

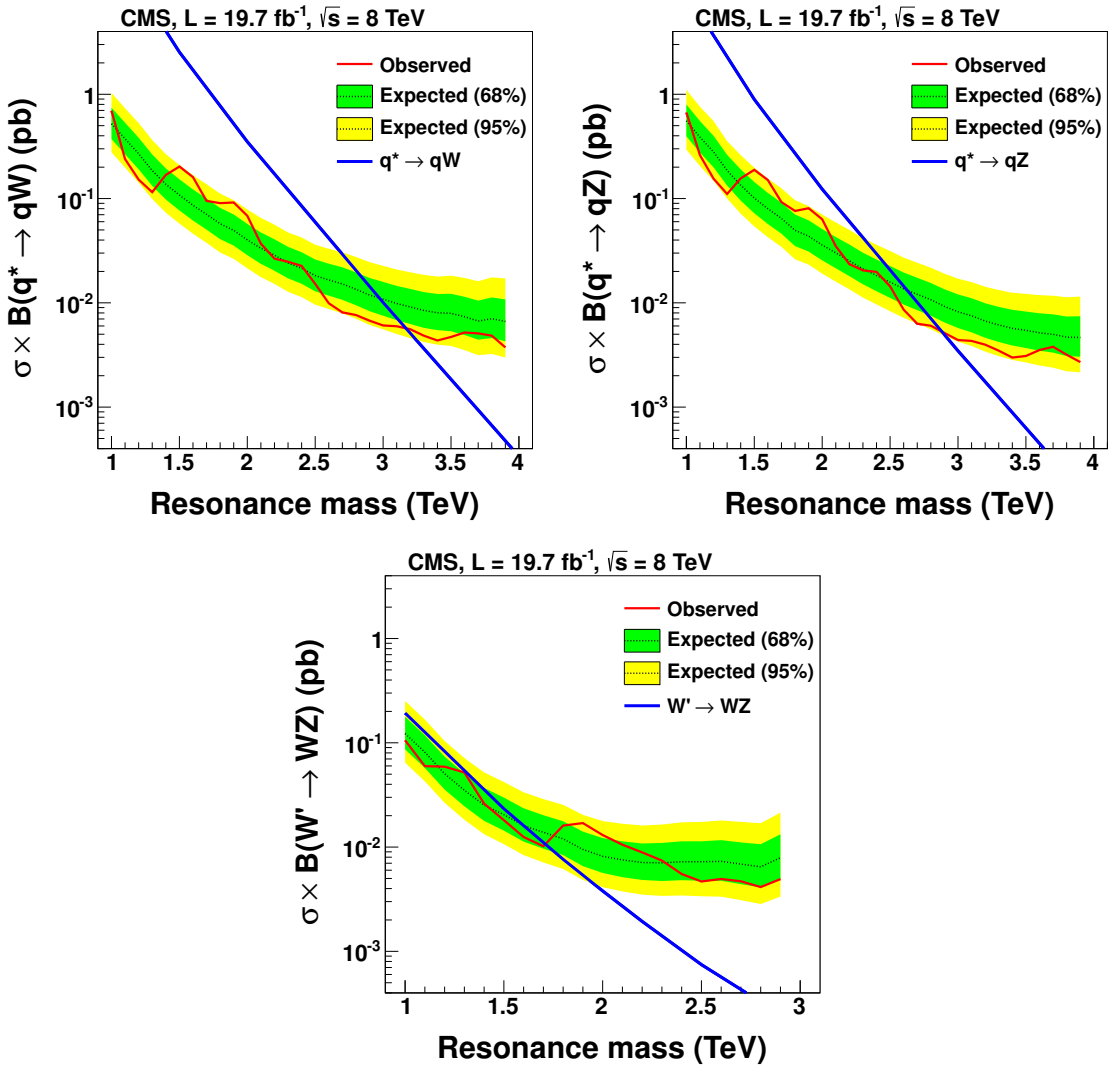

Figure 3.35: Expected and observed 95% CL limits on the production cross section as a function of the resonance mass for (upper left) qW resonances, (upper right) qZ resonances, and (bottom) WZ resonances, compared to their predicted cross sections for the corresponding benchmark models.

computed at 95% CL, with the predicted cross sections for the benchmark models overlaid for comparison. Table 4.8 shows the resulting exclusion ranges on resonant masses. Compared to the previous search in this channel at $\sqrt{s} = 7$ TeV [27], the mass limits on $q* \to q$W and $q* \to q$Z are increased, respectively, by 0.8 and

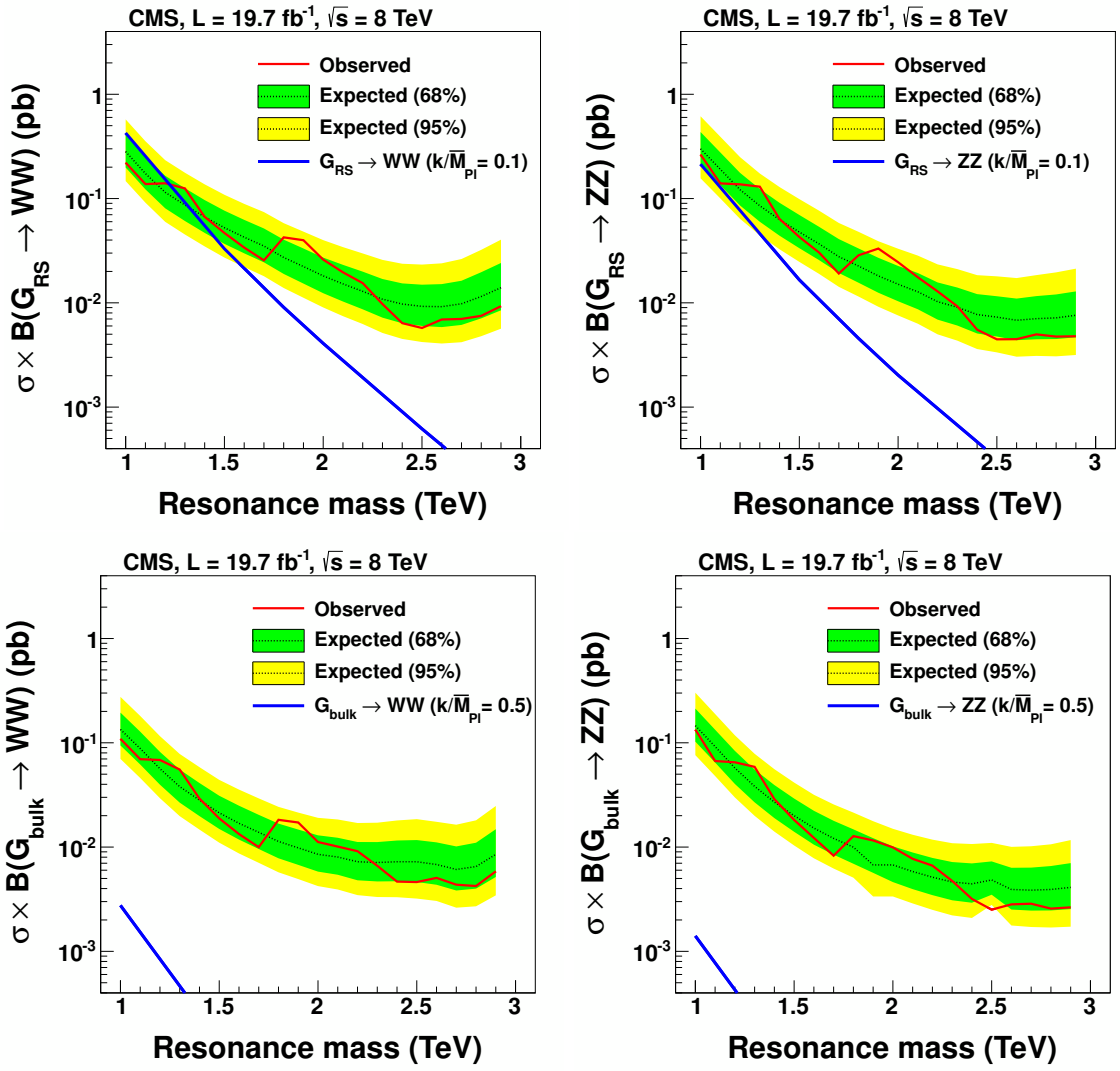

Figure 3.36: Expected and observed 95% CL limits on the production cross section as a function of the resonance mass for (upper left) $G_{\rm RS} \to$ WW resonances, (upper right) $G_{\rm RS} \to$ ZZ resonances, (bottom left) $G_{\rm Bulk} \to$ WW resonances, and (bottom right) $G_{\rm Bulk} \to$ ZZ resonances, compared to the predicted cross sections.

0.7 TeV and for the first time mass limits are set on W$' \to$ WZ and $G_{\rm RS} \to$ WW models. No mass limits are set on $G_{\rm RS} \to$ ZZ, $G_{\rm Bulk} \to$ WW and $G_{\rm Bulk} \to$ ZZ, since the analysis is not sensitive to the small predicted cross sections.

The systematic uncertainties have minor impact on the limits. The largest contri-

butions are 5%, 5%, and 3% from W/Z-tagging efficiency, JES, and JER, respectively. These numbers are obtained by quoting the largest change in the observed exclusion limit on the $G_{\mathrm{RS}} \to \mathrm{WW}$ production cross section, over the entire examined mass range, when the corresponding uncertainties are removed.

Table 3.8: Summary of observed limits on resonance masses at 95% CL and their expected values, assuming a null hypothesis. The analysis is sensitive to resonances heavier than 1 TeV .

Process	Observed	Expected
	excluded mass limit (TeV)	excluded mass limit (TeV)
$q* \to q\mathrm{W}$	3.2	3.0
$q* \to q\mathrm{Z}$	2.9	2.6
$\mathrm{W}' \to \mathrm{WZ}$	1.7	1.6
$G_{\mathrm{RS}} \to \mathrm{WW}$	1.2	1.3

Chapter 4

Search for X $\rightarrow$ WH or ZH at LHC at $\sqrt{s} = 8$ TeV

4.1 Introduction

Several theories of physics beyond the standard model (SM) predict the existence of vector resonances with masses above 1 TeV that decay into a W or Z vector boson (V) and a SM-like Higgs boson (H). Here we present a search for the production of such resonances in proton-proton (pp) collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV . The data sample, corresponding to an integrated luminosity of 19.7fb^{-1}, was collected with the CMS detector at the CERN LHC.

The composite Higgs [78–80] and little Higgs models [81] address the issue of the hierarchy problem and predict many new particles, including additional gauge bosons,

e.g. heavy spin-1 W′ or Z′ bosons. These models can be generalized in the Heavy Vector Triplet (HVT) [82] framework. Of particular interest for this search is the HVT scenario B model, where the branching fraction $\mathcal{B}(W' \to WH)$ and $\mathcal{B}(Z' \to ZH)$ dominate over the corresponding branching fractions to fermions, and are comparable to $\mathcal{B}(W' \to WZ)$ and $\mathcal{B}(Z' \to WW)$. In this scenario, experimental constraints from searches for boson decay channels are more stringent than those from fermion decay channels. Several searches [37,83–86] for $W' \to WZ$ based upon the Extended Gauge Boson (EGB) reference model [18] have excluded resonance masses below 1.7 TeV . Unlike the HVT scenario B model, the EGB model has enhanced fermionic couplings and the mass limit is not directly comparable to this work. Model independent limits on the cross section for the resonant production $\ell\nu + \text{jets}$ [87] can be used to extract resonance mass limits on the the processes $W' \to WZ$ and $Z' \to WW$ of 1.7 TeV and 1.1 TeV, respectively. A search for $Z' \to ZH \to q\bar{q}\tau\tau$ was reported in Ref. [88] and interpreted in the context of HVT scenario model B; however, no resonance mass limit could be set with the sensitivity achieved. Finally, a recent search [89] combining leptonic decays of W and Z bosons, and two b-tagged jets forming a $H \to b\bar{b}$ candidate excluded HVT model A with coupling constant $g_V = 1$ for heavy vector boson masses below $m_{V'^0} < 1360$ GeV and $m_{V'^\pm} < 1470$ GeV.

The signal of interest is a narrow heavy vector resonance V′ decaying into VH, where the V decays to a pair of quarks and the H decays either to a pair of b quarks, or to a pair of W bosons, which further decay into quarks. The H in the HVT framework

does not have properties that are identical to those of a SM Higgs boson. We make the assumption that the state observed by the LHC Collaborations [90, 91] is the same as the one described by the HVT framework and that, in accord with present measurements [92–94], its properties are similar to those of a SM Higgs boson.

In the decay of massive V′ bosons produced in the pp collisions at the LHC, the momenta of the daughter V and H are large enough (>200GeV) that their hadronic decay products are reconstructed as single jets [19]. Because this results in a di-jet topology, traditional analysis techniques relying on resolved jets are no longer applicable. The signal is characterized by a peak in the dijet invariant mass (m_{jj}) distribution over a continuous background from mainly QCD multijet events. The sensitivity to b-quark jets from H decays is enhanced through subjet or jet b tagging [95]. Jets from W/Z → qq′, H → $\mathrm{b\bar{b}}$, and H → WW* → 4q (virtual W denoted with an asterisk) decays are identified with jet substructure techniques [25, 26].

This is the first search for heavy resonances decaying via VH into all-jet final states and it incorporates the first application of jet substructure techniques to identify H → WW* → 4q at a high Lorentz boost.

This analysis proceeds via the following steps:

1. The search is performed in the dijet sample, using the same preselection as the standard search for resonances decaying to dijets [70, 96].

2. We identify events with one W/Z boson jet: a candidate jet originating from merged decaying products of W/Z:

- we require a pruned jet mass cut, and
- an N-subjettiness cut preferring two-prong decays

(This is identical to Chapter 3.)

3. We identify events with a highly boosted Higgs boson:
   - we require a pruned jet mass cut, and
   - two b tagged subjets, or
   - (when there are no two b tagged subjets) a N-subjettiness cut preferring four subjets

   (The H → $b\bar{b}$ tagging is synchronized with our sister analysis, the Radion search to the HH final state [97].)

4. After the full event selection, a potential signal would be characterized as a peak in the dijet invariant mass, on top of a falling background distribution.

5. We model the background distribution with a smoothly falling analytical function. (The functional form is identical to the one used in Chapter 3.)

6. We form the joint likelihood of several dijet distributions of V tagged and H tagged jets. We include both two types of Higgs tags, and also low-purity Higgs and V taggers. The background estimate procedure is the same in all channels – analytical parametrization – but is performed separately for each channel.

7. Finally, we set the limits on the various simplified models for resonances decaying to HV final states.

4.2 Data and Monte Carlo samples

The data sample of proton-proton collisions at $\sqrt{s} = 8$ TeV was collected in 2012 and corresponds to an integrated luminosity of 19.7 fb^{-1}. It is the same as the data studied in Chapter 3 Table 3.1. The dijet data sample is dominated by light flavored and gluon jets, which we denote as the "QCD background".

In the HVT framework, the production cross sections of W′ and Z′ bosons and their decay branching fractions depend on three parameters in addition to the resonance masses: the strength of couplings to quarks (c_{q}), to the H (c_{H}), and on their self-coupling (g_{V}). In the HVT model B, where $g_{\mathrm{V}} = 3$ and $c_{\mathrm{q}} = -c_{\mathrm{H}} = 1$, W′ and Z′ preferentially couple to bosons (W/Z/H), giving rise to diboson final states. This feature reproduces the properties of the W′ and Z′ bosons predicted by the minimal composite Higgs model. In this case, the production cross sections for Z′, W′$^{-}$, and W′$^{+}$ are respectively 165, 87, and 248 fb for a signal of resonance mass $m_{\mathrm{V}'} = 1$ TeV. Their branching fractions to VH and decay width are respectively 51.7%, 50.8%, 50.8% and 35.0, 34.9, 34.9 GeV. The resonances are assumed to be narrow, *i.e.*, with natural widths smaller than the experimental resolution in m_{jj} for masses considered in this analysis.

We consider the W′ and Z′ resonances separately, and report limits for each candidate individually to permit the reinterpretation of our results in different scenarios with different numbers of spin-1 resonances.

Signal events are simulated using the MADGRAPH 5.1.5.11 [98] Monte Carlo

(MC) event generator to generate partons that are then showered with with PYTHIA 6.426 [99] to produce final state particles. These events are then processed through a GEANT4 [65] based simulation of the CMS detector. The MADGRAPH input parameters are provided in Ref. [100] and the H mass is assumed to be 125 GeV . Samples showered with HERWIG++ 2.5.0 [64] are used to evaluate the systematic uncertainty associated with the hadronization. Tune Z2* [68] is used in PYTHIA, while the version 23 tune [64] is used in HERWIG++. The CTEQ6L1 [66] parton distribution functions (PDF) are used for MADGRAPH, PYTHIA, and HERWIG++. Signal events are generated from resonance mass 1.0 to 2.6 TeV in steps of 0.1 TeV. Signals with resonance masses between the generated values are interpolated. Part of the signal samples and their cross sections are listed in Table 4.1.

Process	mass (GeV)	Events	X-sec[pb]
Z' → HZ	1000	20000	8.56E-02
	1500	20000	1.19E-02
	2000	20000	1.93E-03
	2500	20000	3.39E-04
W' → HW	1000	20000	1.71E-01
	1500	20000	2.55E-02
	2000	20000	4.25E-03
	2500	20000	7.31E-04

Table 4.1: Examples of the simulated Monte Carlo samples used in this analysis for process $V' \to VH$. Cross sections are calculated by its production cross sections of V' times its $\mathcal{B}$($W' \to HW$ or $Z' \to HZ$).

4.3 Event selections

The event reconstruction adopt the same algorithm and procedure as Section 3.2 in Chapter 3. For details, please refer to that section.

4.4 The H tagging and W/Z tagging algorithms

In this analysis, we aim to cover as much of the Higgs branching fraction as possible. The Standard Model Higgs with a mass of 125 GeV decays to $b\bar{b}$ with a branching fraction of 57.7%, and to WW* with a branching fraction of 21.4% [101]. Using these two decay modes in a VH search, where WW* specifically decays to four quarks, is the main topic of this note.

The algorithms to identify W/Z, H → $b\bar{b}$ and $H \rightarrow WW^* \rightarrow 4q$ jets are necessarily different, but they use similar jet-level variables: N-subjettiness (described in Section 4.4.1) and jet pruning (Section 4.4.2). The W/Z-tagger is described in Section 4.4.3, and the two H-taggers in Sections 4.4.4 and 4.4.5.

4.4.1 N-subjettiness

The details of N-subjettiness is elaborated before. Please refer to Section 3.3.1 in Chapter 3.

4.4.2 Jet Pruning

Please refer to Section 3.3 in Chapter 3 for the procedure of jet pruning.

Here the result of jet pruning on the CA8 jets is two fold, i.e., the invariant jet

mass reconstruction and subjet identification. In all cases, we use the jet invariant mass computed from the whole (or "fat") pruned jet. This quantity is referred below to as the pruned jet mass. For W/Z tagging, we use pruned jet mass between 70 and 100 GeV . For the identification of Higgs jets, we require the pruned jet mass to lie between 110 and 135 GeV . The distribution of the pruned jet mass of the Higgs candidate jet compared to W/Z and top jets is shown on Figure 4.1.

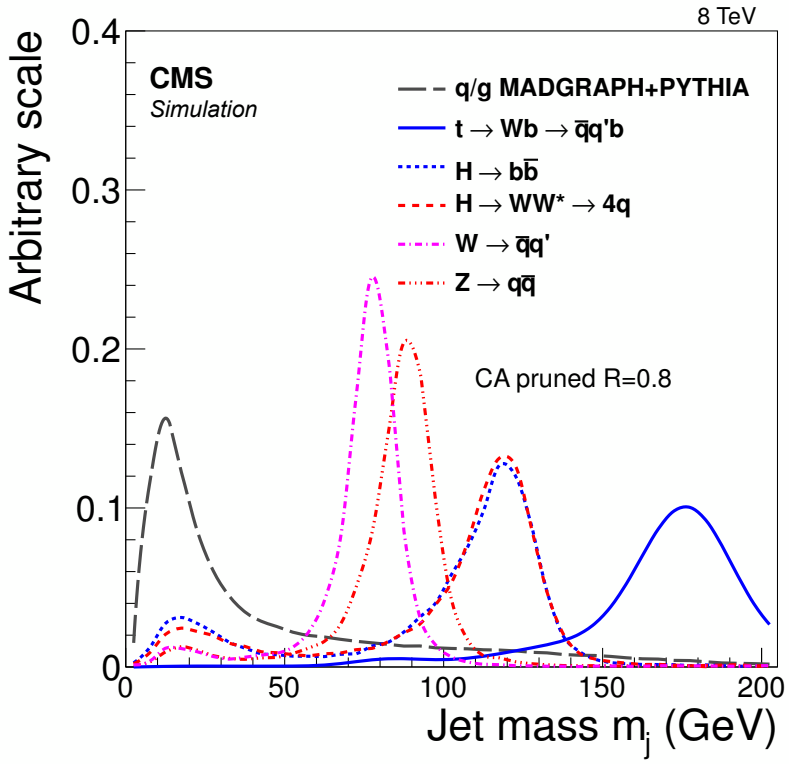

Figure 4.1: Distribution of pruned jet mass in simulation of signal and background processes. All simulated distributions are normalized to 1. The W/Z, H, and top-quark jets are required to match respective generator level particles in the event. The W/Z and H jets are from 1.5 TeV $W' \to WH$ and $Z' \to ZH$ signal samples.

The main role of jet pruning is to allow better delineation of subjets within the

jet. In H → $b\bar{b}$ tagger, the axes of the pruned subjets are used as the basis for b tagging.

4.4.3 W/Z tagging

For the identification of W/Z jets, we employ the same tagging algorithm previously used in Section 3.3 in Chapter 3. The W/Z jets are selected using the following requirements:

- **Pruned jet mass m_{jet}** - Require the total pruned jet mass to satisfy $70 < m_{\text{jet}} < 100$ GeV .
- **N-subjettiness** - We split the events into two categories, "high purity" W/Z jets by requiring $\tau_{21} \leq 0.5$, while $0.5 < \tau_{21} < 0.75$ defines the "low purity" W/Z jets. The thresholds are the same as those in Chapter 3.

The performance of the W/Z tagger has been documented in detail in Reference [26].

4.4.4 H → $b\bar{b}$

To identify Higgs jets arising from the shower and hadronization of two collimated b quarks, we apply b tagging either on the two subjets or the fat CA8 jet, based on the angular separation of the two subjets ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) , which is recommended by Reference [95]. So for $H \to b\bar{b}$ tagging, we use the following selections,

synchronized with the Radion search [97] and the search for HW resonances in the semileptonic channel [102]:

- **Pruned jet mass m_{jet}** - Require the total pruned jet mass to satisfy $110 < m_{\text{jet}} < 135$ GeV .
- **Subjet b-tagging**
  - if ΔR between the two CA8 subjets is bigger than 0.3: *both* subjets must pass the CSV Loose working point.
  - if ΔR between the CA8 subjets is smaller than 0.3: require the *fat* CA8 jet to pass the CSV Loose working point.

4.4.5 $H \rightarrow WW^* \rightarrow 4q$

In this channel, Higgs decays to two W bosons, one real and one virtual. Given that this is effectively a three-body decay $H \rightarrow Wqq$, the jets from the four quarks are not on an even footing – the subjets from the real W are harder, and they also form a W mass. The subjets from the softer two quarks are less well defined.

A naive $H \rightarrow 4q$ tagger would require a fat jet with four subjets. However, a study done using the subjets defined by the CMS Top Tagging algorithm (which reruns the CA8 jet clustering with additional weak pruning [103]) removes $\approx 90\%$ of the signal. Compounded with a decreasing angular separation between Higgs decay products, as a function of the Higgs p_T, at higher resonance masses, e.g., at 2 TeV ,

only 1% of signal passes this selection. The distribution of the number of subjets of the reconstructed Higgs jets in signal MC is shown in Figure 4.2.

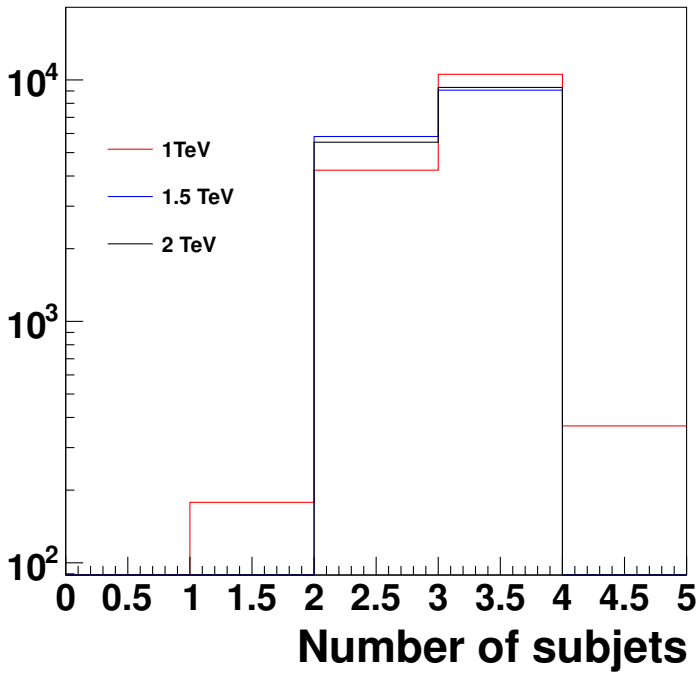

Figure 4.2: Number of subjets of the Higgs candidate jets (selected by a matching to the generator level Higgs particle), in W′ → HW signal MC.

As an alternative, we explore the N-subjettiness, in particular the variables involving τ_4. The ratio $\tau_{42} \equiv \tau_4/\tau_2$ has the best separation between the H → 4q signal and not only QCD background, but also W/Z and top jets. Figures 4.3 and 4.4 show the discriminating power of τ_{42} against $t\bar{t}$ and QCD, for $m_{V'}$ at 1 TeV and 2 TeV resonance masses respectively.

We also explore other combinations of $\tau_{NM} \equiv \tau_N/\tau_M$, which are listed in Ap-

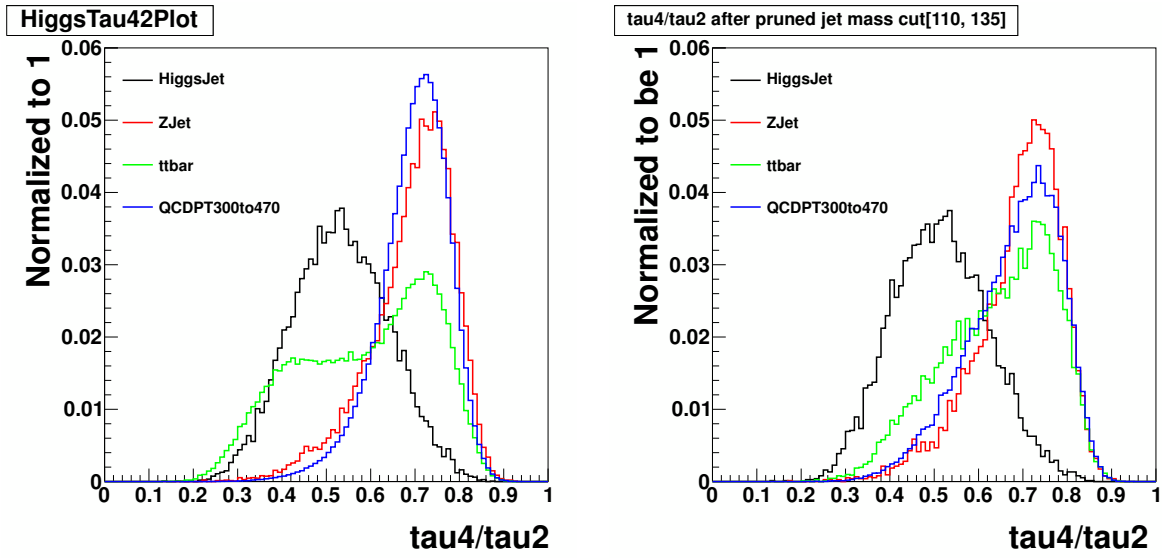

Figure 4.3: Distribution for τ_4/τ_2 in data and in simulations of signal (1.0 TeV) and background events. All simulated distributions are scaled to match the number of events in data, W/Z, matched top and Higgs jets are required to match their generator level particles, respectively.

pendix B.4. The ROC (receiver operating characteristic) curve of for several τ_{NM} cuts (but the same pruned jet mass cut) is shown in Figure 4.5. The signal efficiency is evaluated using Higgs jets in 2 TeV signal MC, and the false positive rate (*i.e.*, mistag rate) is derived from QCDPT300to470 MC sample. From the figure, it is clear that τ_{42} outperforms any other single τ_{NM} variable.

After optimizing the cut on τ_{42} (documented in Section 4.4.5.1 below), the full selection of the H → WW* → 4q tagger is:

- **Pruned jet mass m_{jet}** - We require the total pruned jet mass to satisfy $110 < m_{\text{jet}} < 135$ GeV .

- **N-subjettiness** - We split the events into two categories, "high purity" Higgs

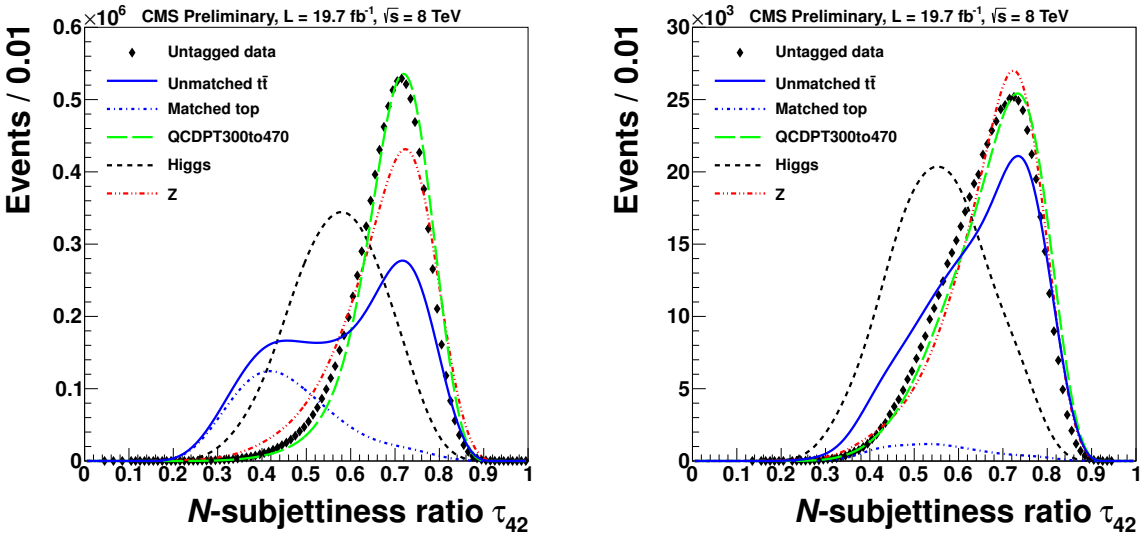

Figure 4.4: Distributions of τ_{42} in data and in simulations of signal (2 TeV) and background events, without applying the pruned jet mass requirement (left) and with the pruned jet mass requirement applied (right). Matched top-quark, W/Z, and H_{WW} jets are required to be consistent with their generator level particles, respectively. All simulated distributions are scaled to the number of events in data, except that matched top-quark background is scaled to the fraction of unmatched $t\bar{t}$ events times the number of data events.

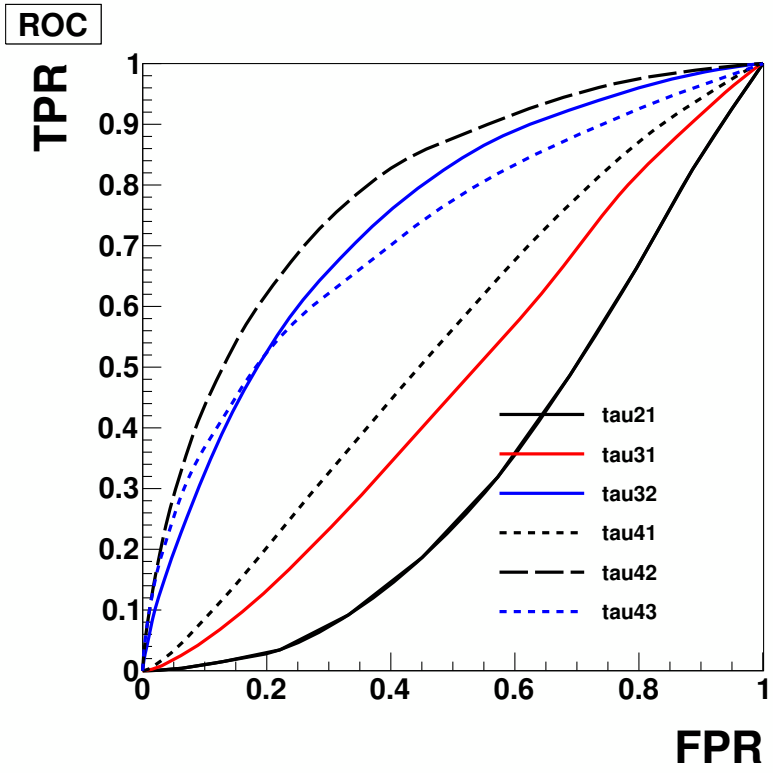

Figure 4.5: ROC curves for different τ_{NM} after the cut on the pruned jet mass. The false positive rate (FPR) is obtained from QCDPT300to470 and the true positive rate (TPR) from Higgs jets in 2 TeV signal MC sample. Using τ_{42} to select Higgs jets outperforms all other τ_{NM} variables.

jets by requiring $\tau_{42} \leq 0.55$, while $0.55 < \tau_{42} < 0.65$ defines the "low purity" Higgs jets.

4.4.5.1 Optimization of the τ_4/τ_2 threshold

Having selected τ_{42} as the discriminating variable, we next optimize its upper value. In this study, the jet mass is confined within $[110, 135]$ GeV. We use the limit setting method (described in Section 3.9) and evaluate the expected limits of several signal resonance masses at different τ_{42} working points. These expected limits are presented in Table 4.2. Given our focus on the resonance masses above 1500 GeV , we choose to cut on $\tau_{42} < 0.55$. In the following analysis, to compensate the signal efficiency loss at higher resonance mass, we introduce an additional categories for $H \to \mathrm{WW}^* \to 4\mathrm{q}$ tagger as $0.55 < \tau_{42} < 0.65$. This is chosen from back-of-envelope calculation based on Figures 4.3 and 4.4, since this category provides very limited sensitivity.

Table 4.2: Upper limits (in units of 0.01 pb) for high purity HW and HZ signals at different resonance masses with different τ_{42} working points.

HW / τ_{42}	0.45	0.5	0.55	0.6
1000	4.14	4.09	4.46	4.91
1500	0.97	0.88	0.86	0.91
2000	0.89	0.64	0.51	0.47
2500	1.36	0.82	0.53	0.40
HZ/ τ_{42}				
1000	4.31	4.36	4.63	5.05
1500	0.98	0.89	0.86	0.90
2000	0.70	0.55	0.42	0.39
2500	0.96	0.61	0.41	0.32

4.5 Trigger

We use the same triggers as in Section 3.5. Events are selected if one of the following triggers has fired: HLT_HT750, HLT_PFHT650, HLT_PFNoPUHT650, HLT_FatDiPFJetMass750_DR1p1_Deta1p5. Figures 4.7, 4.8 and 4.9 show the trigger efficiency. The trigger efficiency has been measured with respect to a lower-thrseshold, but prescaled, HLT_HT550 trigger. The trigger is 99% effiecient above 890 GeV for the untagged, HbbVqq-tagged , and HwwVqq-tagged data.

Figure 4.6 shows the turn-on curve of the reference trigger on the signal MC. The 1.0 TeV signals are used here in Figure 4.6, and the plot shows that the HwwVqq and HbbVqq signals are fully efficient for HLT_HT550 trigger, which is not prescaled in MC. So other signals, having resonance mass bigger than 1.0 TeV , will surely be fully efficient for the triggers.

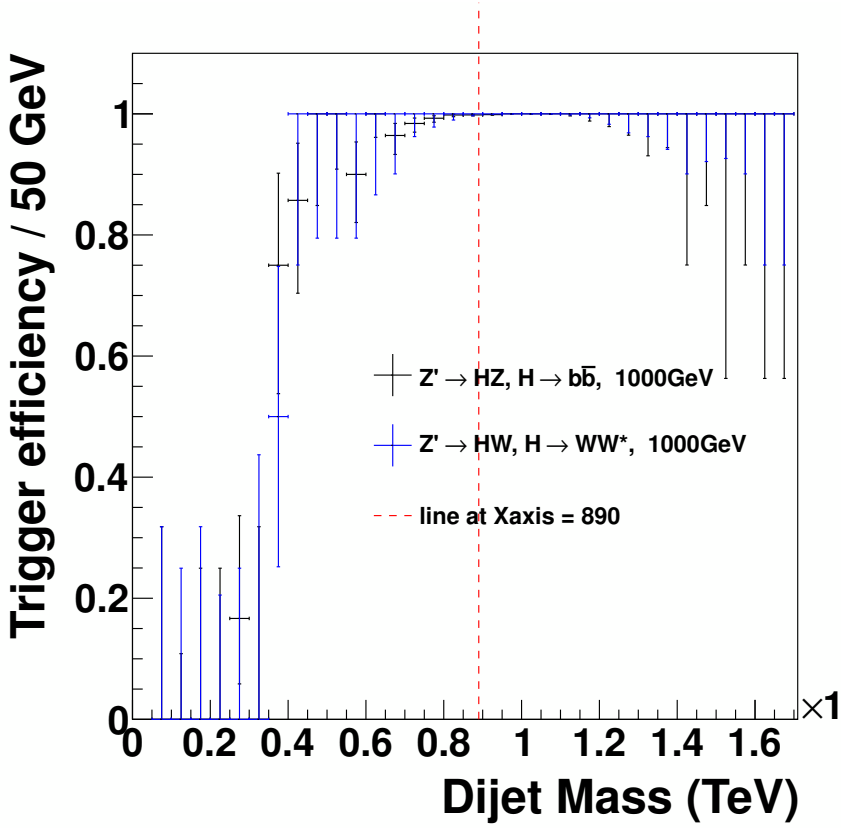

Figure 4.6: Reference trigger efficiency of signal MC.

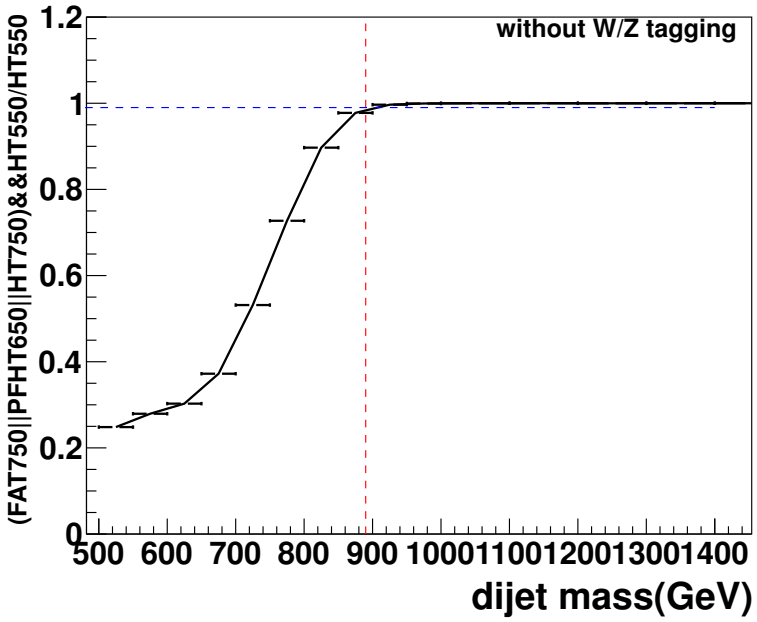

Figure 4.7: trigger efficiency for untagged data of fat_750||hlt_pf(nopu)ht650||hlt_ht750 measured using data collected by lower threshold h_t550 trigger. the dashed red line is drawn at m_{jj} equal 890 GeV, the blue line is at efficiency at 99%.

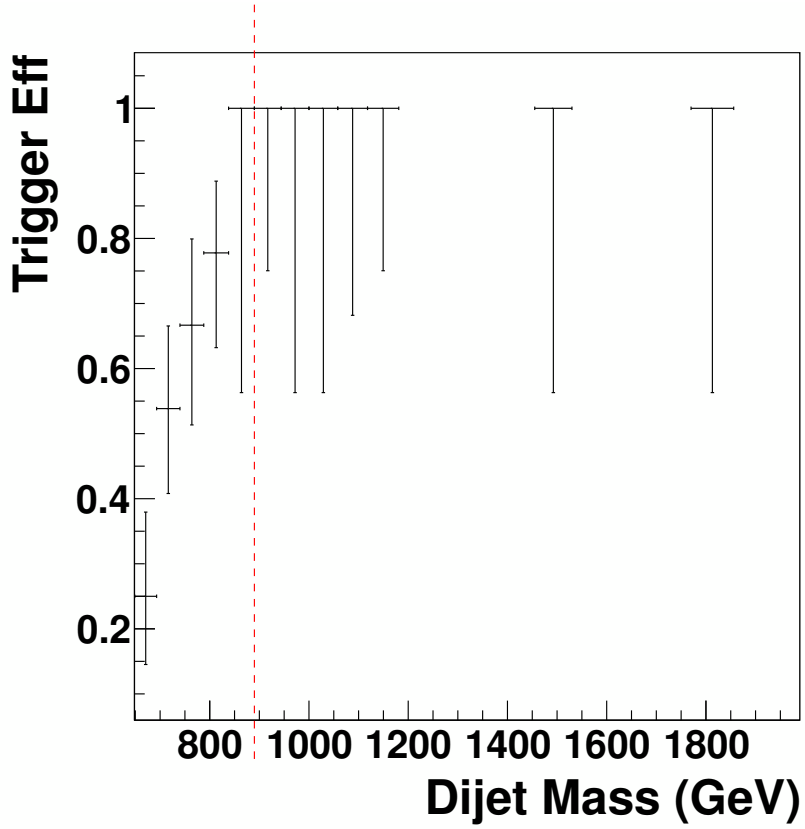

Figure 4.8: Trigger efficiency for HbbVqq tagged data of FAT_750‖HLT_PF(NoPU)HT650‖HLT_HT750 measured using data collected by lower threshold H_T550 trigger. The dashed red line is drawn at m_{jj} equal 890 GeV.

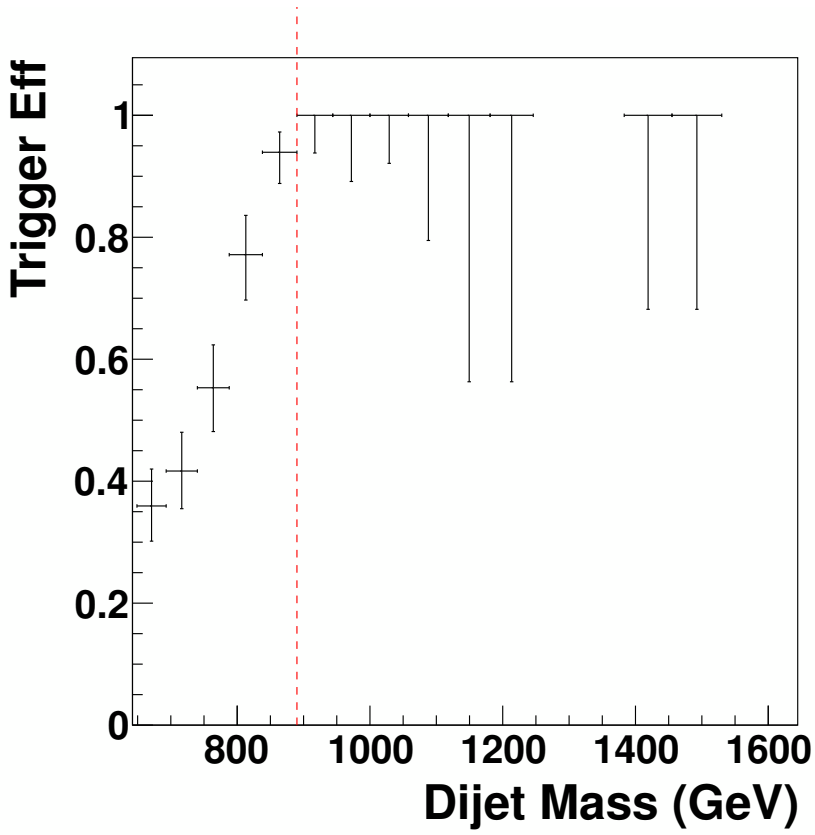

Figure 4.9: Trigger efficiency for HqqqqVqq tagged data of FAT_750‖HLT_PF(NoPU)HT650‖HLT_HT750 measured using data collected by lower threshold H_T550 trigger. The dashed red line is drawn at m_{jj} equal 890 GeV.

4.6 The signals

We search for several models of heavy resonances decaying to a W or Z boson on one side, and a Higgs on the other, in which both bosons decay to quarks producing merged jet. This analysis is focused on two channels:

- $H \to \mathrm{b\bar{b}}, \mathrm{W/Z} \to qq'$, and
- $H \to \mathrm{WW^*} \to \mathrm{4q}, \mathrm{W/Z} \to qq'$

As previously discussed, we use one V-tagging and two Higgs tagging algorithms to identify such events. After subdividing the events according to high purity and low purity tags, we end up with five distinct categories, as shown in Table 4.4.

In this section, we discuss various issues related to the evaluation of the signal efficiency.

4.6.1 Cross-talk between the Higgs decay channels

In order to combine events from all categories into a single joint likelihood, the categories must be mutually exclusive. However, a cross-talk between the Higgs channels is nevertheless possible: for example, $H \to \mathrm{b\bar{b}}$ tagger can identify other two-prong Higgs decay modes like $\mathrm{H} \to \mathrm{gg}$, $H \to \tau\tau$, or $\mathrm{H} \to \mathrm{c\bar{c}}$, although this kind of 'false positive' tag happens only rarely (the efficiency is $\lesssim 7$ %). Similarly, events from two-prong Higgs decay channels can also pass the τ_{42} cut in the $H \to \mathrm{WW^*} \to \mathrm{4q}$ selection. In this case, the channel $H \to \mathrm{b\bar{b}}$, because of its large branching ratio,

contributes a non-negligible number of events to the sample of 4q tags. This effect is illustrated by Fig. 4.10, where it can be seen that most of the low-τ_{42} tail of the $H \to b\bar{b}$ curve will be below the cut value of 0.55.

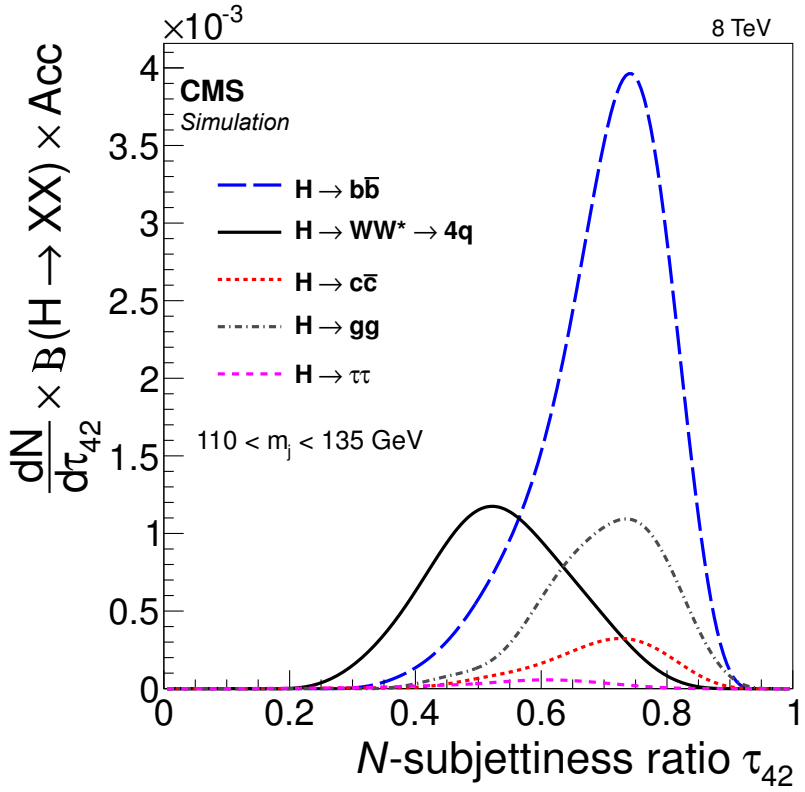

Figure 4.10: Comparison of τ_{42} distributions for signal events failing the $H \to b\bar{b}$ requirement. These events are from the $H \to WW^* \to 4q$, $H \to b\bar{b}$, $H \to gg$, $H \to c\bar{c}$, and $H \to \tau\tau$ channels. The H jets are from a 1.5 TeV resonance decaying to VH. All curves are normalized to the product of the corresponding branching fraction and acceptance.

Table 4.3 provides an overview of the cross-talk between the various channels. The Higgs branching ratios correspond to the Higgs mass of 125 GeV . For $H \to WW^* \to 4q$, the branching ratio of the hadronic decay of (real) W boson is already included, so that the final state is four quarks. The table is normalized to 100,000 standard model

Higgs bosons, and the numbers in the table show the number of Higgs decays that pass the tagger for each channel, with the branching ratio taken into account. For example, let us consider H → $c\bar{c}$ channel. At the Z′ resonance mass of 1.5 TeV, out of 100,000×3.0% H → $c\bar{c}$ decays, 121 of them are tagged by the $H \to b\bar{b}$ tagger and 88 of them pass $H \to WW^* \to 4q$ tagger but fail $H \to b\bar{b}$ tagger. For H → ZZ decays, we take its tagging efficiency the same as $H \to WW^* \to 4q$ signals. So the number of H → ZZ to pass $H \to b\bar{b}$ and $H \to WW^* \to 4q$ tagger is estimated by efficiency of $H \to WW^* \to 4q$ signal times $\mathcal{B}(H \to ZZ)\times\mathcal{B}(Z \to qq)\times\mathcal{B}(Z \to qq)$ divided by $\mathcal{B}(H \to WW)\times\mathcal{B}(W \to qq)\times\mathcal{B}(W \to qq)$.

Since the $H \to b\bar{b}$ tagger has significantly lower background than $H \to WW^* \to 4q$, it takes precedence in selecting events: we first identify the events that pass the $H \to b\bar{b}$ tagger, and only if they fail, we test them for the presence of the $H \to WW^* \to 4q$ tag.

The effect of the $H \to b\bar{b}$ tagger veto on the $H \to WW^* \to 4q$ tagged dijet mass distribution background (data) is shown in Appendix B.5.

Table 4.3: Number of Higgs jets falls into two exclusive categories, assuming we have 100,000 SM Higgs (125 GeV) decays to all channels.

	Branching ratio (%)	Pass $H \to b\bar{b}$	Fail $H \to b\bar{b}$, pass $H \to WW^* \to 4q$
1.5 TeV			
$H \to b\bar{b}$	57.70	11444	755
$H \to WW^* \to 4q$	9.94	228	1916
$H \to ZZ^* \to 4q$	1.30	29	250
$H \to c\bar{c}$	3.00	121	88
$H \to \tau\tau$	6.30	12	57
$H \to gg$	10.00	69	174
2.0 TeV			
$H \to b\bar{b}$	57.70	13816	551
$H \to WW^* \to 4q$	9.94	449	1435
$H \to ZZ^* \to 4q$	1.30	58	187
$H \to c\bar{c}$	3.00	228	99
$H \to \tau\tau$	6.30	42	74
$H \to gg$	10.00	157	262

4.6.2 Summary of Higgs and W/Z tagging categories

The W or Z jets from the signal are selected by the V-tagger, and the Higgs candidates are selected by an OR of the two Higgs taggers, $H \to b\bar{b}$ and $H \to WW^* \to 4q$. Both V-tagger and $H \to WW^* \to 4q$ taggers have high-purity and low-purity categories. The latter are added to increase the sensitivity of the analysis at high resonance masses, where the QCD background is low, and a higher signal efficiency is at the premium.

We first identify the events that pass the $H \to b\bar{b}$ tagger, and only if they fail, we test them for the presence of the $H \to WW^* \to 4q$ tag. Thus we arrive at the final division of events into five mutually exclusive categories listed in Table 4.4. For the $H \to WW^* \to 4q, W/Z \to qq'$ channel, we drop the low-purity Higgs and low-purity V-tagging category, because it adds only a negligible sensitivity.

The events from the $H \to b\bar{b}, W/Z \to qq'$ signals could contribute to all the five categories, due to its large branching ratio. The $H \to WW^* \to 4q, W/Z \to qq'$ signal events contribute only in events that fail $H \to b\bar{b}$ but pass $H \to WW^* \to 4q$ tagger; their contribution to $H \to b\bar{b}$ tagged sample is negligible. The contributions from other Higgs decay modes to all these five categories is tiny compared to $H \to b\bar{b}, W/Z \to qq'$ and $H \to WW^* \to 4q, W/Z \to qq'$ yields. We will not specifically study them, but include them as systematic uncertainties.

Table 4.4: Summary of event categories and their nomenclature used in this search. The jet mass cut is $70 < m_j < 100$ GeV for the V tag and $110 < m_j < 135$ GeV for the H tag.

Categories	V tag	H tag
$V^{HP}H_{bb}$	$\tau_{21} \leq 0.5$	b tag
$V^{LP}H_{bb}$	$0.5 < \tau_{21} < 0.75$	b tag
$V^{HP}H^{HP}_{WW}$	$\tau_{21} \leq 0.5$	$\tau_{42} \leq 0.55$
$V^{LP}H^{HP}_{WW}$	$0.5 < \tau_{21} < 0.75$	$\tau_{42} \leq 0.55$
$V^{HP}H^{LP}_{WW}$	$\tau_{21} \leq 0.5$	$0.55 < \tau_{42} < 0.65$

4.6.3 Signal acceptance and efficiencies

To enable the results to be applied to other models of similar final states, we utilize simulations to derive the geometrical acceptances and the selection efficiencies, presented separately in Figures 4.11, 4.12, and 4.13. The global efficiency is approximated by the product of acceptances and the W/Z and H tagging efficiency, restricted to final states where the W/Z and H bosons decay hadronically. A matching of the generated W, Z, and H bosons, and their reconstructed single jets is required within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.5$ as a part of the acceptances. The products of acceptances and the W/Z and H tagging efficiency, ignoring leptonic decays and the correlations between detector acceptance and W/Z or H tagging, agree to better than 10% with the full event simulation. In the interpretations reported in this search, the global efficiency is estimated from the full simulation of signal events, without apply-

ing the matching requirement. In this way, the correlations between the acceptance and W/Z and H tagging efficiency are properly taken into account. However, when interpreting this search in terms of W/Z and H tagging efficiency for an arbitrary model an additional uncertainty of 10% should be folded in.

The acceptance, shown in Fig. 4.11 as a function of the dijet resonance mass for several signals, takes into account the angular acceptance ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$) and the matching of the W, Z, and H bosons with their reconstructed single jets.

The expected tag probabilities of the W, Z, and H selection criteria for signal and data events in different event categories are shown in Figs. 4.12 and 4.13, as a function of m_{jj}. The W/Z and $H \to \mathrm{WW}^* \to 4\mathrm{q}$ tagging efficiencies for signal events in the HP categories drop at high $\mathrm{p_T}$, primarily because the τ_{21} and τ_{42} distributions are $\mathrm{p_T}$-dependent.

The MC modelling of V-tag efficiency is validated using high-$\mathrm{p_T}$ $\mathrm{W} \to \mathrm{q}'\bar{\mathrm{q}}$ decays selected from a data sample enriched in semileptonic $\mathrm{t\bar{t}}$ events [26]. Scale factors of 0.86 ± 0.07 and 1.39 ± 0.75 are applied to the MC events in the HP and LP V tag categories, respectively, to match the tagging efficiencies in the top pair data. The decay of $H \to \mathrm{WW}^* \to 4\mathrm{q}$ produces a hard W jet accompanied by two soft jets from the off-shell W boson. As the $H \to \mathrm{WW}^* \to 4\mathrm{q}$ tagger is also based on the N-subjettiness variables, and the measured ratio τ_{42}/τ_{21} is well modelled by QCD simulation, it is reasonable to assume that the mismodelling of the shower by PYTHIA is similar to that in the case of V tagging. The $H \to \mathrm{WW}^* \to 4\mathrm{q}$ tagging

efficiency scale factors are extrapolated using the same technique as for V tagging for both the HP and LP categories, respectively, with additional systematic uncertainties, which are discussed in Section 4.8. The resolution for the m_{jj} reconstruction is in the range $5 - 10\%$ for all the five categories.

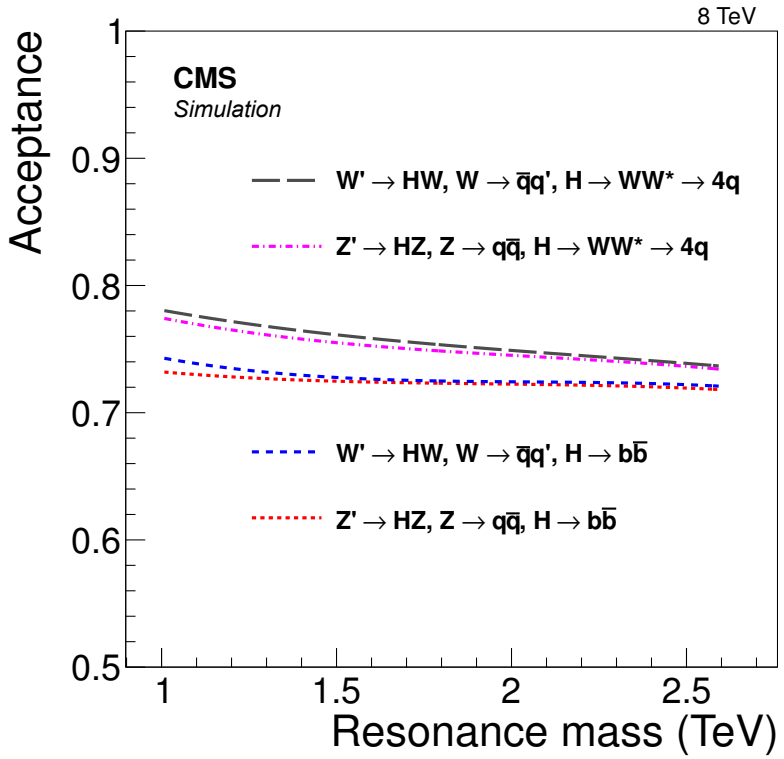

Figure 4.11: The fraction of simulated signal events for hadronically decaying W/Z and H bosons, reconstructed as two jets, that pass the geometrical acceptance criteria ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$), shown as a function of the resonance mass.

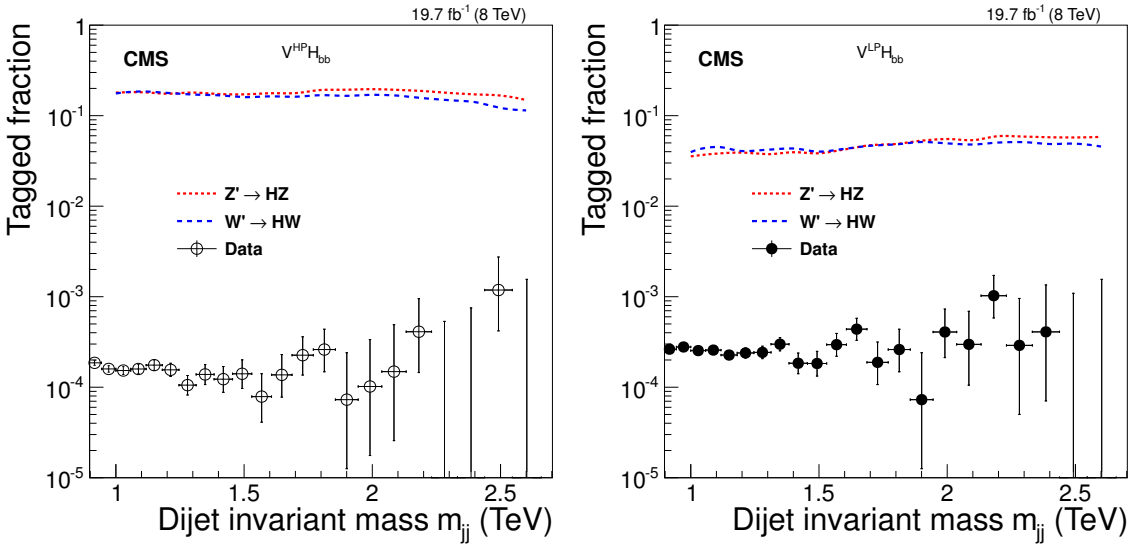

Figure 4.12: Tagged fractions in $H \to b\bar{b}, W/Z \to qq'$ signal channels and data as a function of dijet invariant mass, for categories of $V^{HP}H_{bb}$ (left) and $V^{LP}H_{bb}$ (right). Horizontal bars through the data points indicate the bin width.

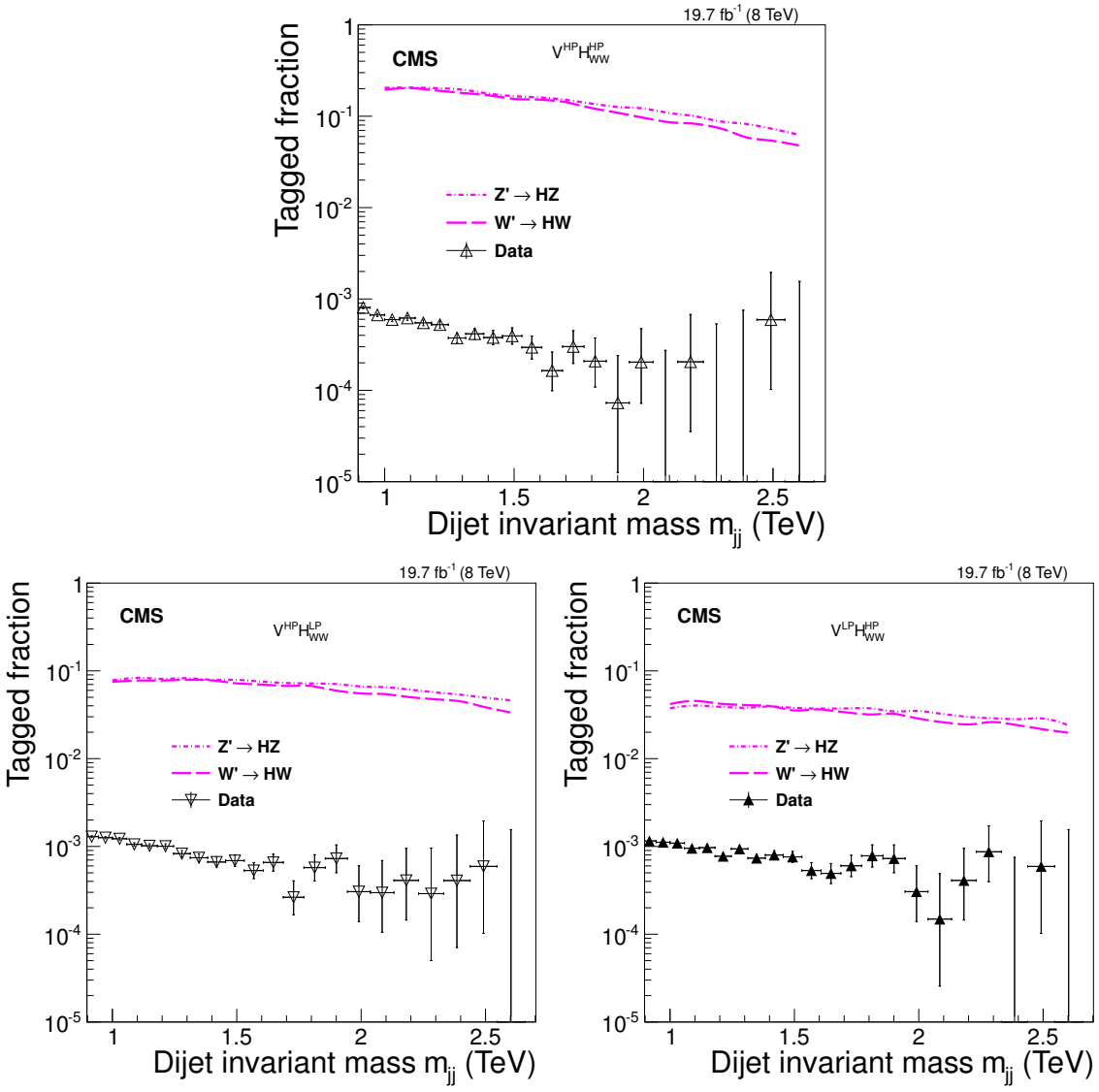

Figure 4.13: Tagged fractions in $H \to \mathrm{WW}^* \to 4\mathrm{q}, \mathrm{W/Z} \to qq'$ signal channels and data as a function of dijet invariant mass, for categories of $\mathrm{V^{HP}H^{HP}_{WW}}$ (top), $\mathrm{V^{HP}H^{LP}_{WW}}$ (bottom left) and $\mathrm{V^{LP}H^{HP}_{WW}}$ (bottom right). Horizontal bars through the data points indicate the bin width.

4.7 Resonance Search in the dijet mass spectrum

The resolution for the m_{jj} reconstruction is in the range $5 - 10\%$ for all the five categories. The background from multijet events is modelled by a smoothly falling distribution for each event category, given by the empirical probability density function

$$P_D(m_{\mathrm{jj}}) = \frac{P_0(1 - m_{\mathrm{jj}}/\sqrt{s})^{P_1}}{(m_{\mathrm{jj}}/\sqrt{s})^{P_2}} \ . \tag{4.1}$$

Each event category has separate normalization P_0 and shape parameters P_1 and P_2. This parameterization was deployed successfully in a number of searches based on dijet mass spectra [70]. A Fisher F-test [104] is used to check that no additional parameters are needed to model the individual background distributions, compared with the four-parameter function used in [70]. We have also tested an alternative function $P_E(m_{\mathrm{jj}}) = P_0/(m_{\mathrm{jj}}/\sqrt{s} + P_1)^{P_2}$, and found it less favored by the F-test. The use of the alternative function in the analysis produces negligible changes in the final result and therefore, no systematic uncertainty is associated with this choice.

We search for a peak on top of the falling background spectrum by means of a binned maximum likelihood fit to the data. The binned likelihood is given by

$$\mathcal{L} = \prod_i \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}, \tag{4.2}$$

where $\lambda_i = \mu N_i(S) + N_i(B)$, μ is a scale factor for the signal, $N_i(S)$ is the number of events expected from the signal, and $N_i(B)$ is the number expected from multijet background. The variable n_i quantifies the number of observed events in the i^{th} m_{jj} bin. The number of background events $N_i(B)$ is described by the functional form of Eq. (4.1). While maximizing the likelihood, μ as well as the parameters of the background function are unconstrained and left floating. For presentational purposes, a binning according to m_{jj} resolution is used in this paper. However, the likelihood is calculated in bins of 1 GeV in m_{jj}, approximating an unbinned analysis, while keeping it computationally manageable.

Figures 4.14 and 4.15 show the m_{jj} distributions in data, binned according to m_{jj} resolution. The solid curves represent the results of the maximum likelihood fit to the data, fixing the number of expected signal events to zero, while the bottom panels show the corresponding pull distributions, quantifying the agreement between the background-only hypothesis and the data. The expected distributions of $H \to \text{b}\overline{\text{b}}, \text{W/Z} \to qq'$ and $H \to \text{WW}^* \to \text{4q}, \text{W/Z} \to qq'$ signals at 1.0, 1.5 and 2.0 TeV in each category, scaled to their corresponding cross sections are given by the dashed and dash-dotted curves. The resonance masses in VH_{bb} channels are slightly lower than those of the VH_{WW} channels because of missing neutrinos in b-hadron decays and partial misreconstruction of two-pronged $\text{H} \to \text{b}\overline{\text{b}}$ decays.

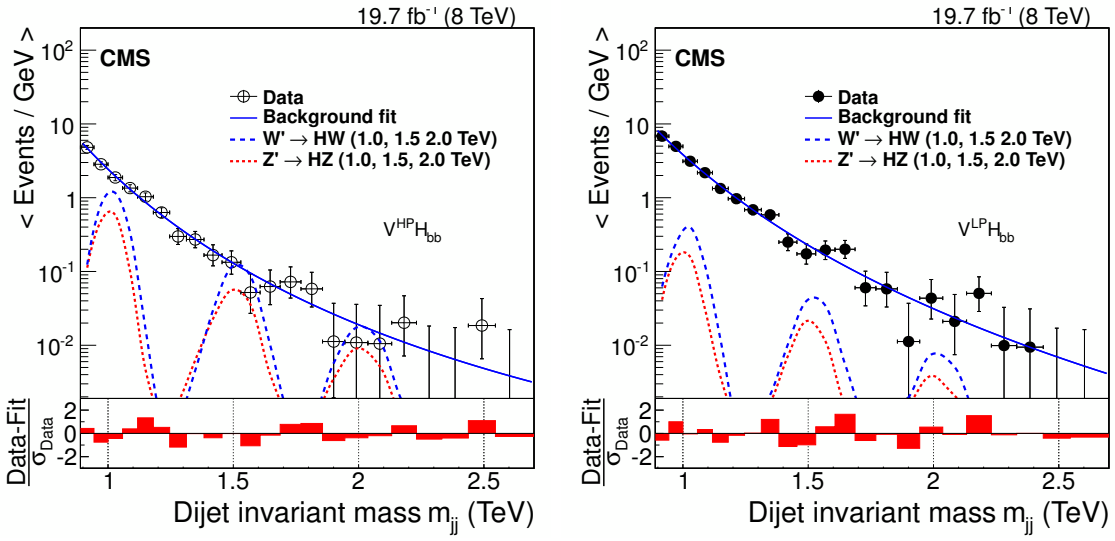

Figure 4.14: Distributions in m_{jj} are shown for $\mathrm{V^{HP}H_{bb}}$ category (left), $\mathrm{V^{LP}H_{bb}}$ category (right). The solid curves represent the results of fitting Eq. (4.1) to the data. The distributions for $H \to \mathrm{b\bar{b}}, \mathrm{W/Z} \to qq'$contributions, scaled to their corresponding cross sections, are given by the dashed curves. The vertical axis displays the number of events per bin, divided by the bin width. Horizontal bars through the data points indicate the bin width. The corresponding pull distributions $\frac{\mathrm{Data-Fit}}{\sigma_{\mathrm{Data}}}$, where σ_{Data} represents the statistical uncertainty in the data in a bin in m_{jj}, are shown below each m_{jj} plot.

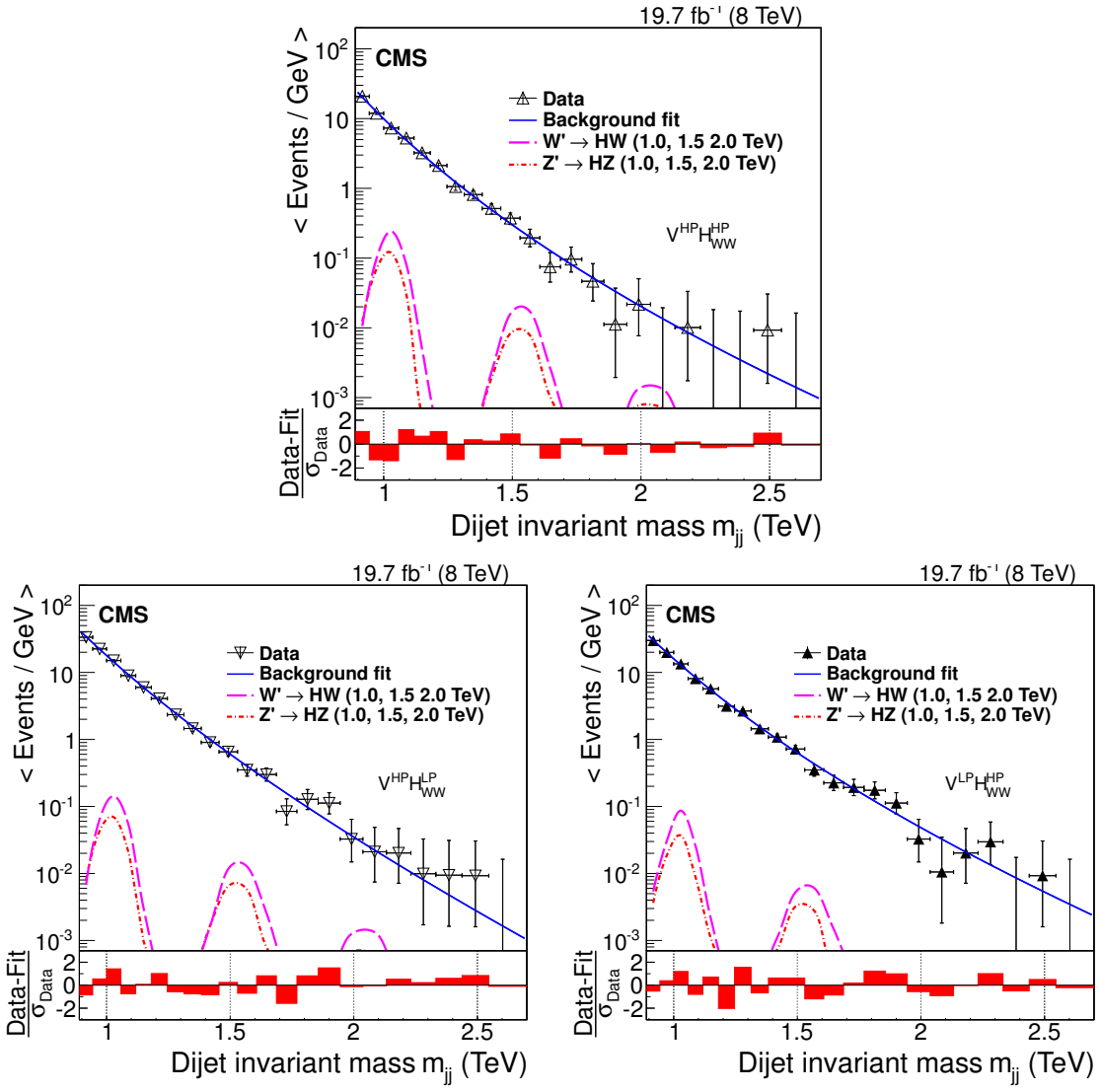

Figure 4.15: Distributions in m_{jj} are shown for $\mathrm{V^{HP}H^{HP}_{WW}}$ (top), $\mathrm{V^{HP}H^{LP}_{WW}}$ (bottom left), and $\mathrm{V^{LP}H^{HP}_{WW}}$ (bottom right). The solid curves represent the results of fitting Eq. (4.1) to the data. The distributions for $H \to \mathrm{WW^*} \to \mathrm{4q}, \mathrm{W/Z} \to qq'$contributions, scaled to their corresponding cross sections, are given by the dashed and dash-dotted curves. The vertical axis displays the number of events per bin, divided by the bin width. Horizontal bars through the data points indicate the bin width. The corresponding pull distributions $\frac{\mathrm{Data-Fit}}{\sigma_{\mathrm{Data}}}$, where σ_{Data} represents the statistical uncertainty in the data in a bin in m_{jj}, are shown below each m_{jj} plot.

4.8 Systematic uncertainties

The largest contributions to the systematic uncertainty are associated with the modelling of the signal, namely: the efficiencies of W/Z, H, and b tagging; the choice of PDF; the jet energy scale (JES); the jet energy resolution (JER); the pileup corrections; the cross-talk between different signal contributions; and the integrated luminosity.

The uncertainty in the efficiency for W/Z-tagging is estimated using a control sample enriched with $t\bar{t}$ events described in Ref. [26]. Uncertainties of 7.5% and 54% in the respective scale factors for HP and LP V tag include contributions from control-sample statistical uncertainties, and the uncertainties in the JES and JER for pruned jets [105]. To extrapolate to higher jet p_T, an estimation of V tagging efficiency varying as a function of $p_T$for two different showering and hadronization models using PYTHIA 6 and HERWIG++, shows that the differences are within 4% (12%) for the HP (LP) V-tagging [26]. We extrapolate the $H \to WW^* \to 4q$ tagging efficiency scale factor in the same way as the W/Z-tagging efficiency, with an additional systematic uncertainty based on the difference between PYTHIA 6 and HERWIG++ in modelling $H \to WW^* \to 4q$ decay. This is evaluated to be ≈7 % for the HP and LP H tag. The uncertainty from the pruned jet mass requirement in the $H \to WW^* \to 4q$ search is already included in the extrapolated scale factor uncertainty of the V-tag.

The uncertainty in the efficiency of $H \to b\bar{b}$ tagging can be separated into two

categories: the efficiency related to the b tagging and the efficiency related to the pruned H mass tag. The first is obtained by varying the b tagging scale factors within the associated uncertainties [95] and amounts to 15%. The second is assumed to be similar to the mass selection efficiency of W jets estimated in Ref. [26], additionally accounting for the difference in fragmentation of light quarks and b quarks, which amounts to 2.6% per jet.

Because of the rejection of charged particles not originating from the primary vertex, and the application of pruning, the dependence of the W/Z and H tagging efficiencies on pileup is weak and the uncertainty in the modelling of the pileup distribution is $\leq 1.5\%$ per jet.

In this analysis, we only consider $H \rightarrow \mathrm{b\bar{b}}$ and $H \rightarrow \mathrm{WW}^* \rightarrow 4\mathrm{q}$ decays. Other H decay channels that pass H taggers are viewed as nuisance signals, and a corresponding cross-talk systematic uncertainty is assigned. We evaluate this uncertainty as a ratio of expected nuisance signal events with respect to the total expected signal events, taking into account the branching fractions, acceptances and tagging efficiencies. The contamination from cross-talk is estimated to be $2 - 7\%$ in the $\mathrm{VH_{bb}}$ categories, and $18 - 24\%$ in the $\mathrm{VH_{WW}}$ categories, and we take the maximum as the uncertainty. The analysis is potentially 7% (24%) more sensitive than quoted, but since it is not clear how well the efficiency for the nuisance signals is understood, they are neglected, yielding a conservative limit on new physics. When the $\mathrm{VH_{bb}}$ and $\mathrm{VH_{WW}}$ categories are combined together, the 24% uncertainty becomes a small effect, based

on a quantitative measure of sensitivity suggested in Ref. [106] :

$$P = \frac{\mathcal{B}(\mathrm{H} \to \mathrm{XX}) \times \epsilon_S}{1 + \sqrt{N_B}} \tag{4.3}$$

where $\mathcal{B}(\mathrm{H} \to \mathrm{XX})$ is the branching fraction for the H decay channel, ϵ_S is the signal tagging efficiency, and N_B is the corresponding background yield. The values of P for each channel are shown in Table 4.5.

Table 4.5: Summary of the values P for a Z' signal at 1.5 TeV resonance mass and the corresponding background yield in all five categories.

Signal/Categories	$\mathrm{V^{HP}H_{bb}}$	$\mathrm{V^{LP}H_{bb}}$	$\mathrm{V^{HP}H^{HP}_{WW}}$	$\mathrm{V^{HP}H^{LP}_{WW}}$	$\mathrm{V^{LP}H^{HP}_{WW}}$
$H \to \mathrm{b\bar{b}}, \mathrm{Z} \to \mathrm{q\bar{q}}$	2.3×10^{-2}	4.8×10^{-3}	1.0×10^{-3}	1.6×10^{-3}	3.9×10^{-4}
$H \to \mathrm{WW^*} \to 4\mathrm{q}, \mathrm{Z} \to \mathrm{q\bar{q}}$	5.6×10^{-4}	≈ 0	2.6×10^{-3}	9.8×10^{-4}	4.5×10^{-4}

The JES has an uncertainty of 1–2% [54,71], and its $\mathrm{p_T}$and η dependence is propagated to the reconstructed value of m_{jj}, yielding an uncertainty of 1%, independent of the resonance mass. The impact of this uncertainty on the calculated limits is estimated by changing the dijet mass in the analysis within its uncertainty. The JER is known to a precision of 10%, and its non-Gaussian features observed in data are well described by the CMS simulation [54]. The effect of the JER uncertainty on the limits is estimated by changing the reconstructed resonance width within its uncertainty. The integrated luminosity has an uncertainty of 2.6% [72], which is also taken into

account in the analysis. The uncertainty related to the PDF used to model the signal acceptance is estimated from the CT10 [107], MSTW08 [108], and NNPDF21 [109] PDF sets. The envelope of the upward and downward variations of the estimated acceptance for the three sets is assigned as uncertainty [110] and found to be 5 – 15% in the resonance mass range of interest. A summary of all systematic uncertainties is given in Table 4.6 and 4.7. Among these uncertainties, the JES and JER are applied as shape uncertainties, while others are applied as uncertainty in the event yield.

Table 4.6: Systematic uncertainties common to all categories.

Source	HP uncertainties (%)	LP uncertainties (%)
JES	1	1
JER	10	10
Pileup	≤ 3.0	≤ 3.0
PDF	5–15	5–15
Integrated luminosity	2.6	2.6
W-tagging	7.5	54
W tag $p_T$dependence	4	12

Table 4.7: Systematic uncertainties(%) for X $\rightarrow$ VH signals, in which $H \rightarrow b\bar{b}$ and $H \rightarrow WW^* \rightarrow 4q$. Numbers in parentheses represent the uncertainty for the corresponding LP category. If LP has the same uncertainty as HP, only the HP uncertainty is presented here.

Source/Final State	$H \rightarrow b\bar{b}$		$H \rightarrow WW^* \rightarrow 4q$
	VH_{bb}	VH_{WW}	VH_{WW}
$H \rightarrow b\bar{b}$ mass scale	2.6	-	-
H(4q)-tagging	-	7.5 (54)	7.5 (54)
H(4q)-tag τ_{42} extrapolation	-	7	7
Cross-talk	7	24	24
b-tagging	≤ 15	≤ 15	-

4.9 Results

The asymptotic approximation [74] of the LHC CL_s criterion [75, 76] is used to set upper limits on the cross section for resonance production. The dominant sources of systematic uncertainties are treated as nuisance parameters associated with log-normal priors in those variables. For a given value of the signal cross section, the nuisance parameters are fixed to the values that maximize the likelihood, a method referred to as profiling. The dependence of the likelihood on parameters used to describe the background in Eq. (4.1) is treated in the same manner, and no additional systematic uncertainty is assigned to the parameterization of the background.

Events from the 5 categories of Table 4.4 are combined into a common likelihood, with the uncertainties of the HP and LP H tag (V tag) efficiencies considered to be anticorrelated between HP and LP tagging because events failing the HP τ_{42} (τ_{21}) selection migrate to the LP category and the fraction of events failing both HP and LP requirements is small compared to the HP and LP events. The branching fractions of $H \to WW^* \to 4q$ and $H \to b\bar{b}$ decays are taken as fixed values in joint likelihood. The remaining systematic uncertainties in the signal are fully correlated across all channels. The variables describing the background uncertainties are treated as uncorrelated. Figure 4.16 shows the observed and background-only expected upper limits on the production cross sections for Z' and W', including both $H \to b\bar{b}$ and $H \to WW^* \to 4q$ decays, computed at 95% confidence level (CL), with the predicted cross sections for the benchmark models overlaid for comparison. In the HVT model

scenario B, W' and Z' are degenerate in resonance mass, thus we compute the limit on their combined cross section under this hypothesis, shown in Fig. 4.17. Table 4.8 shows the exclusion ranges on resonance masses.

Table 4.8: Summary of observed lower limits on resonance masses at 95% CL and their expected values, assuming a null hypothesis. The analysis is sensitive to resonances heavier than 1 TeV .

Process	Observed	Expected
	lower mass limit (TeV)	lower mass limit (TeV)
$W' \to HW$	[1.0, 1.6]	1.7
$Z' \to HZ$	[1.0, 1.1], [1.3, 1.5]	1.3
$V' \to VH$	[1.0, 1.7]	1.9

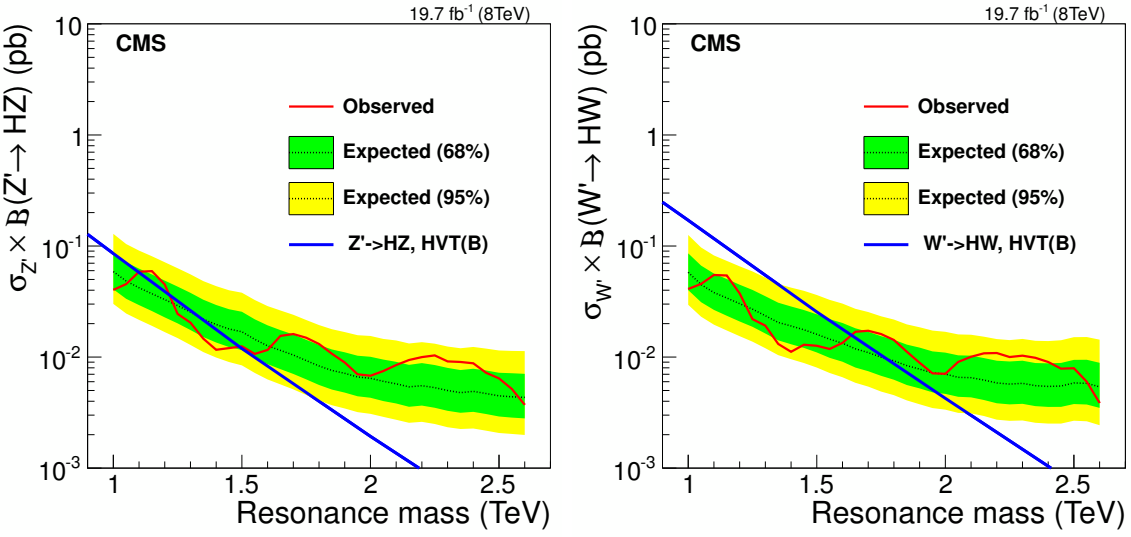

Figure 4.16: Expected and observed upper limits on the production cross sections for $Z' \to HZ$ (left) and $W' \to HW$ (right), including all five decay categories. Branching fractions of H and V decays have been taken into account. The theoretical predictions of the HVT model scenario B are also shown.

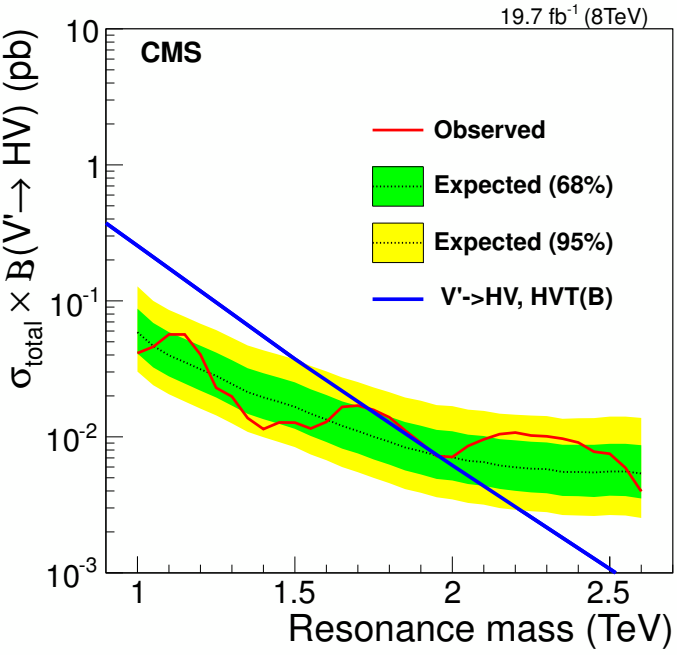

Figure 4.17: Expected and observed upper limits on the production cross section for $V' \to VH$, obtained by combining W' and Z' channels together. Branching fractions of H and V decays have been taken into account. The theoretical prediction of the HVT model scenario B is also shown.

Chapter 5

Conclusions and outlook

Currently all experimental studies in particle physics are in agreement with a set of theories called the Standard Model (SM). However, from the cosmological and other observations, as well as the internal consistency checks, the SM is found to be an incomplete theory. Thus it is believed to be an effective theory of a more fundamental theory. There are many models predicting physics beyond-SM, but so far none of them have been corroborated by experimental results.

This thesis describes several searches for the signatures of models of beyond-SM physics. A common feature of these searches is that their experimental signature is a pair of energetic jets (back-to-back in detector), one or both of which are characterized by the internal jet structure ("substructure").

The jet substructure arises when the W/Z (V) or Higgs (H) boson is at a large Lorentz boost and the hadronization products from its decay ($\mathrm{W/Z} \to \mathrm{qq}'$, $H \to \mathrm{b\bar{b}}$,

or H $\rightarrow$ WW* $\rightarrow$ qqqq) merge into a single jet. Thus the jet from W/Z or Higgs decay has several cores of energy. In contrast, the main source of background – the QCD production – formed by the hadronization of a quark or a gluon, does not possess the jet substructure.

The analysis of the substructure is based on several aspects sensitive to the difference between signal and background:

- pruned jet mass (removing soft and wide-angle radiations) : in the process of jet pruning, the QCD jets lose a lot of mass, whereas the bulk of the mass of W/Z/H jets is mostly preserved.
- N-subjettiness (Equation 3.1) is small if the jet is N-prong. Different ratios of N-subjettiness variables are chosen and optimized for different signals.
- subjet b tagging: for $H \rightarrow b\bar{b}$, we search for b hadron decays in subjets.

5.1 Conclusions for X $\rightarrow$ qV or VV analysis

An inclusive sample of multijet events corresponding to an integrated luminosity of 19.7 fb^{-1}, collected in pp collisions at $\sqrt{s} = 8$ TeV with the CMS detector, is used to measure the W/Z-tagged dijet mass spectrum for the two leading jets, produced within the pseudorapidity range $|\eta| < 2.5$ with a separation in pseudorapidity of $|\Delta\eta| < 1.3$. The generic multijet background is suppressed using jet-substructure tagging techniques that identify vector bosons decaying into $q\bar{q}'$ pairs merged into

a single jet. In particular, the invariant mass of pruned jets and the N-subjettiness ratio τ_{21} of each jet are used to reduce the initially overwhelming multijet background. The remaining background is estimated through a fit to smooth analytic functions.

With no evidence for a peak on top of the smoothly falling background, lower limits are set at the 95% confidence level on masses of excited quark resonances decaying into qW and qZ at 3.2 and 2.9 TeV, respectively. Randall–Sundrum gravitons G_{RS} decaying into WW are excluded up to 1.2 TeV, and W$'$ bosons decaying into WZ, for masses less than 1.7 TeV.

For the first time mass limits are set on $\mathrm{W}' \to \mathrm{WZ}$ and $G_{\mathrm{RS}} \to \mathrm{WW}$ in the all-jets final state. A model with a "bulk" graviton G_{Bulk} that decays into WW or ZZ bosons is also studied, but no mass limits could be set due to the small predicted cross sections.

5.2 Conclusions for $\mathrm{X} \to \mathrm{VH}$ analysis

A search for a massive resonance decaying into a standard model-like Higgs boson and a W or Z boson is presented. A data sample corresponding to an integrated luminosity of 19.7 fb^{-1}collected in proton-proton collisions at $\sqrt{s} = 8$ TeV with the CMS detector has been used to measure the W/Z and Higgs boson-tagged dijet mass spectra using the two highest $\mathrm{p_T}$ jets within the pseudorapidity range $|\eta| < 2.5$ and with pseudorapidity separation $|\Delta\eta| < 1.3$. The QCD background is suppressed

using jet substructure tagging techniques, which identify boosted bosons decaying into hadrons. In particular, the mass of pruned jets and the N-subjettiness ratios τ_{21} and τ_{42}, as well as b tagging applied to the subjets of the Higgs boson jet, are used to discriminate against the otherwise overwhelming QCD background. The remaining QCD background is estimated from a fit to the dijet mass distributions using a smooth function.

We have searched for the signal as a peak on top of the smoothly falling QCD background. No significant signal is observed. In the HVT model B, a Z′ is excluded in resonance mass intervals [1.0, 1.1] and [1.3, 1.5] TeV, while a W′ is excluded in the interval [1.0, 1.6] TeV. A mass degenerate W′ plus Z′ particle is excluded in the interval [1.0, 1.7] TeV.

This is the first search for heavy resonances decaying into a Higgs boson and a vector boson (W/Z) resulting in a hadronic final state, as well as the first application of jet substructure techniques to identify $H \to \mathrm{WW}^* \to 4\mathrm{q}$ decays of the Higgs boson at high Lorentz boost.

5.3 Outlook

The most recent ATLAS search for a $\mathrm{X} \to \mathrm{VV}$ resonance shows an evidence of about 3σ excess at the dijet invariant mass of 2 TeV in Figure 5.1. Similar effects were also observed in our own VV analysis, but with a smaller excess of 1.3 sigma at

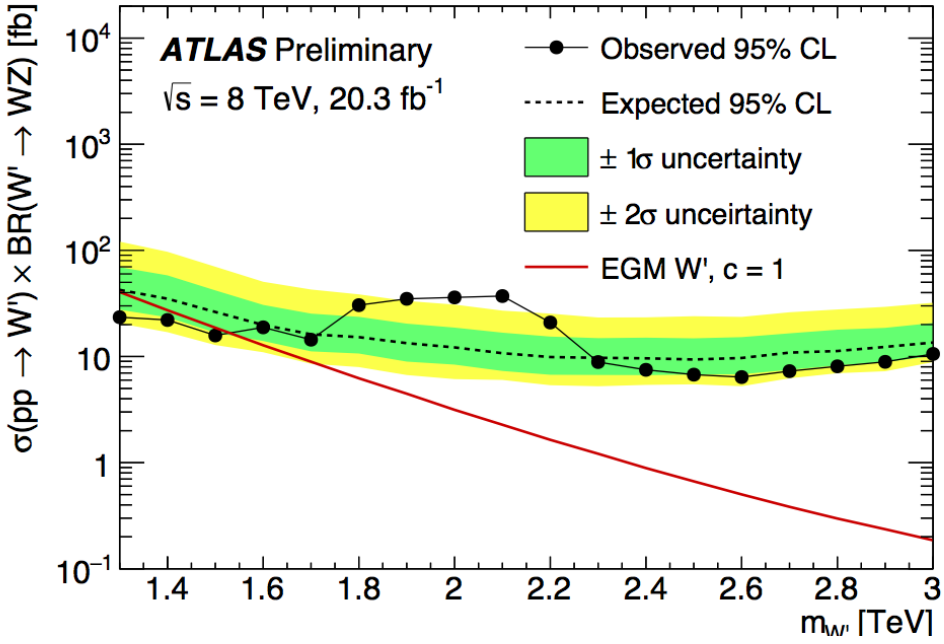

Figure 5.1: ATLAS results on the VV channel analysis.

1.9 TeV shown in Figure 3.36.

At the end of Run I, there are hints, but no statistically significant proof of new physics at ∼2 TeV scale. In 2015, the center-of-mass collision energy at the LHC will reach 13 TeV, with a high luminosity. For $X \to VV$ or $X \to HV$, the cross section is expected to go up ∼30–70 times, but for QCD only by ∼2–3 times compared to now. With more events, we will verify the bump as of just a statistical fluctuation or a discovery of beyond-SM physics.

The discovery of the Higgs boson was a tremendous success of the hard work of thousands of physicists and engineers for the past 30 years. However, it is not the end, for about 96% of the universe (dark matter, dark energy) is beyond the SM and our current understanding. We hope that we are at the beginning of a new era in particle physics, where the data from the LHC can be mined in the years to come and used to flesh out the outlines of the theory that underlies the SM.

Appendix A

Appendices for Chapter 3

A.1 MC generator parameters for the signal models

$q* \to W/Z + jet$ in Pythia6 with Tune Z2:

```
processParmeters = cms.vsting(
        'MSEL=0 ! (D=1) 0 to select full user control',
        'MSTP(6)=1 ! excited quarks',
        'MSUB(147)=1 ! qg->d*',
        'MSUB(148)=1 ! qg->u*',
        'PMAS(343,1)=${mass} ! mass of d*',
        'PMAS(344,1)=${mass} ! mass of u*',
```

```
        'RTCM(41)=${scale} ! Lambda = mass',
        'RTCM(43)=1.0 ! f',
        'RTCM(44)=1.0 ! fp',
        'RTCM(45)=1.0 ! fs',
        '4000001:ALLOFF',
        '4000001:ONIFMATCH 1 23 ! qW=1 23, qZ=2 24',
        '4000002:ALLOFF',
        '4000002:ONIFMATCH 2 23 ! qW=2 23, qZ=1 24',
)
```

$G_{RS} \to WW/ZZ$ in Pythia6 with Tune Z2:

```
processParameters = cms.vstring(
        'PMAS(347,1)=${mass} ! mass of RS Graviton',
        'PARP(50)=${kmpl} ! 0.54 == c=0.1 (k/M_PL=0.1)',
        'MSEL=0 ! (D=1) 0 to select full user control',
        'MSUB(391)=1 ! q qbar -> G*',
        'MSUB(392)=1 ! g g -> G*',
        '5000039:ALLOFF ! Turn off all decays of G*',
        '5000039:ONIFANY 24 ! Turn on the decays WW=24, ZZ=23',
)
```

$G_{RS} \to WW/ZZ$ in Herwig++:

```
   configFiles = cms.vstring('RS.model'),
   parameterSets = cms.vstring(
      'cm7TeV',
      'pdfCTEQ6L1',
      'productionParameters',
      'basicSetup',
      'setParticlesStableForDetector',
   ),
   productionParameters = cms.vstring(
'cd /Herwig/NewPhysics',
'insert ResConstructor:Incoming 0 /Herwig/Particles/g',
'insert ResConstructor:Incoming 1 /Herwig/Particles/u',
'insert ResConstructor:Incoming 2 /Herwig/Particles/ubar',
'insert ResConstructor:Incoming 3 /Herwig/Particles/d',
'insert ResConstructor:Incoming 4 /Herwig/Particles/dbar',
'insert ResConstructor:Intermediates 0 /Herwig/Particles/Graviton',
'insert ResConstructor:Outgoing 0 /Herwig/Particles/W+', #Z0
'set RS/Model:Lambda_pi 10000*GeV',
'set /Herwig/Particles/Graviton:NominalMass ${scale}*GeV',
   )
```

$W' \to WZ$ in Pythia6 with Tune Z2:

```
processParameters = cms.vstring(
        'PMAS(34,1)=\${mass} ! mass of Wprime',
        'MSEL=0 ! (D=1) 0 to select full user control',
        'MSUB(142)=1 ! qq->Wprime',
        'MDME(311,1)=0 ! Wprime->dubar',
     'MDME(312,1)=0 ! Wprime->dcbar',
        'MDME(313,1)=0 ! Wprime->dtbar',
        'MDME(315,1)=0 ! Wprime->subar',
 'MDME(316,1)=0 ! Wprime->scbar',
        'MDME(317,1)=0 ! Wprime->stbar',
        'MDME(319,1)=0 ! Wprime->bubar',
  'MDME(320,1)=0 ! Wprime->bcbar',
        'MDME(321,1)=0 ! Wprime->btbar',
        'MDME(327,1)=0 ! Wprime->enu',
   'MDME(328,1)=0 ! Wprime->munu',
        'MDME(329,1)=0 ! Wprime->taunu',
        'MDME(331,1)=1 ! Wprime->WZ',
        'MDME(332,1)=0 ! Wprime->Wgamma',
        'MDME(333,1)=0 ! Wprime->Wh0',
)
```

A.1.1 Signal shape at high resonance mass

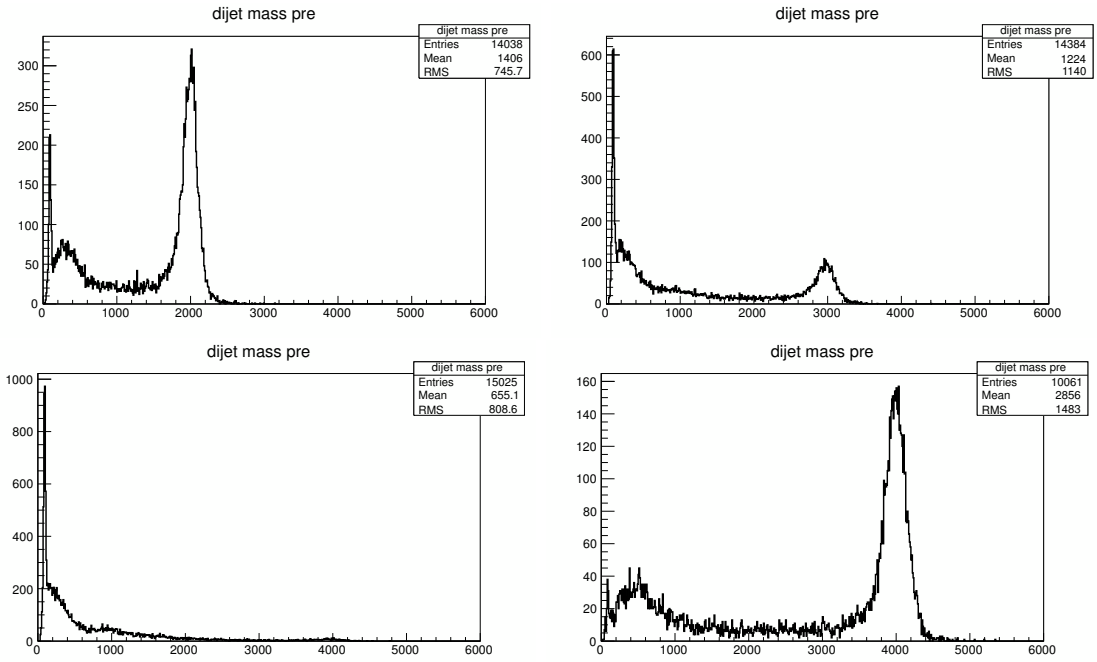

Figure A.1: Comparsion for signal dijet mass distribution for $W' \to WZ$ at 2.0 TeV(top left), 3.0 TeV(top right), 4.0 TeV(bottom left). Plot on the right bottom is the dijet mass distribution for $G_{RS} \to WW$ at 4.0 TeV.

A.2 Event displays

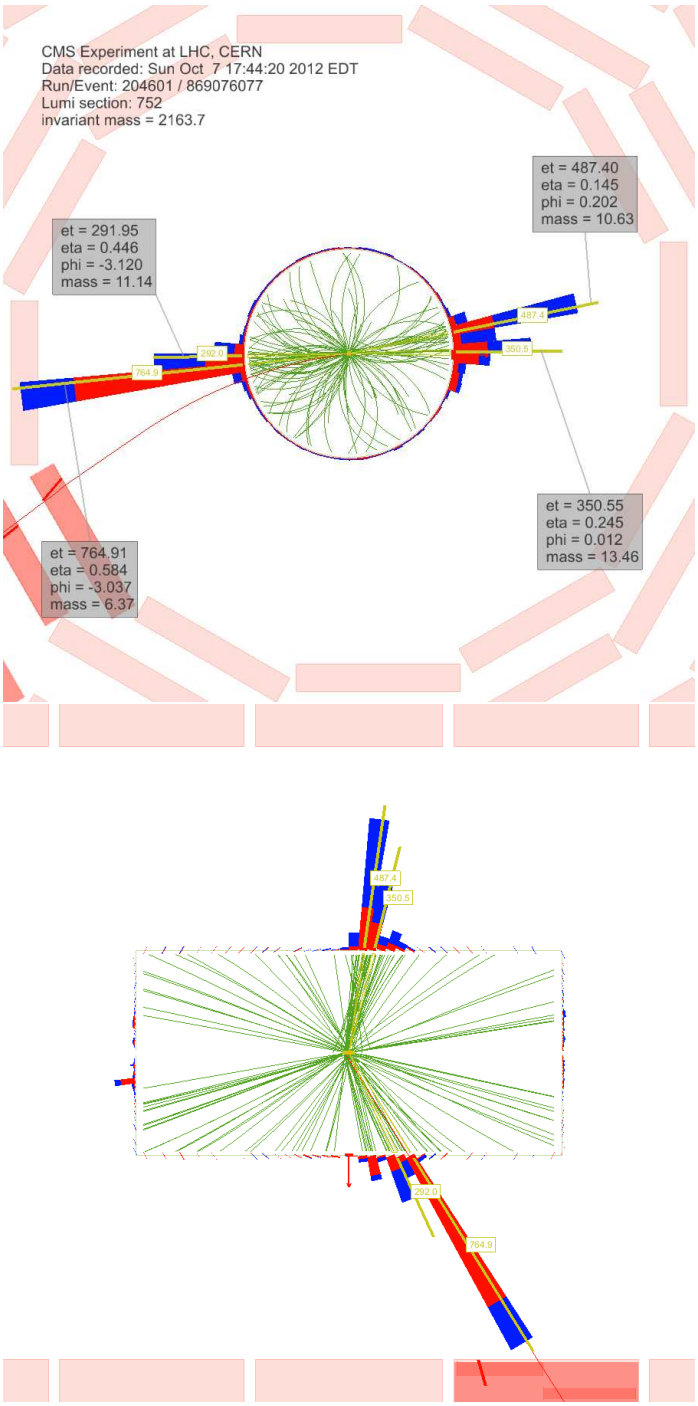

Figure A.2: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

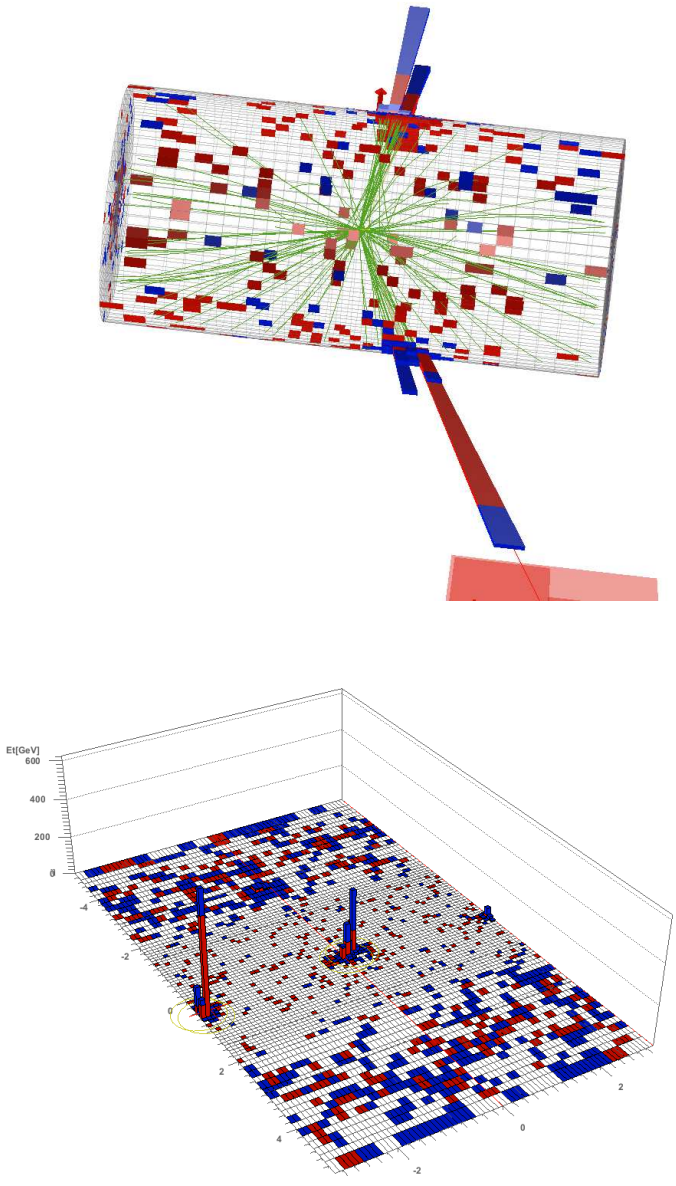

Figure A.3: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

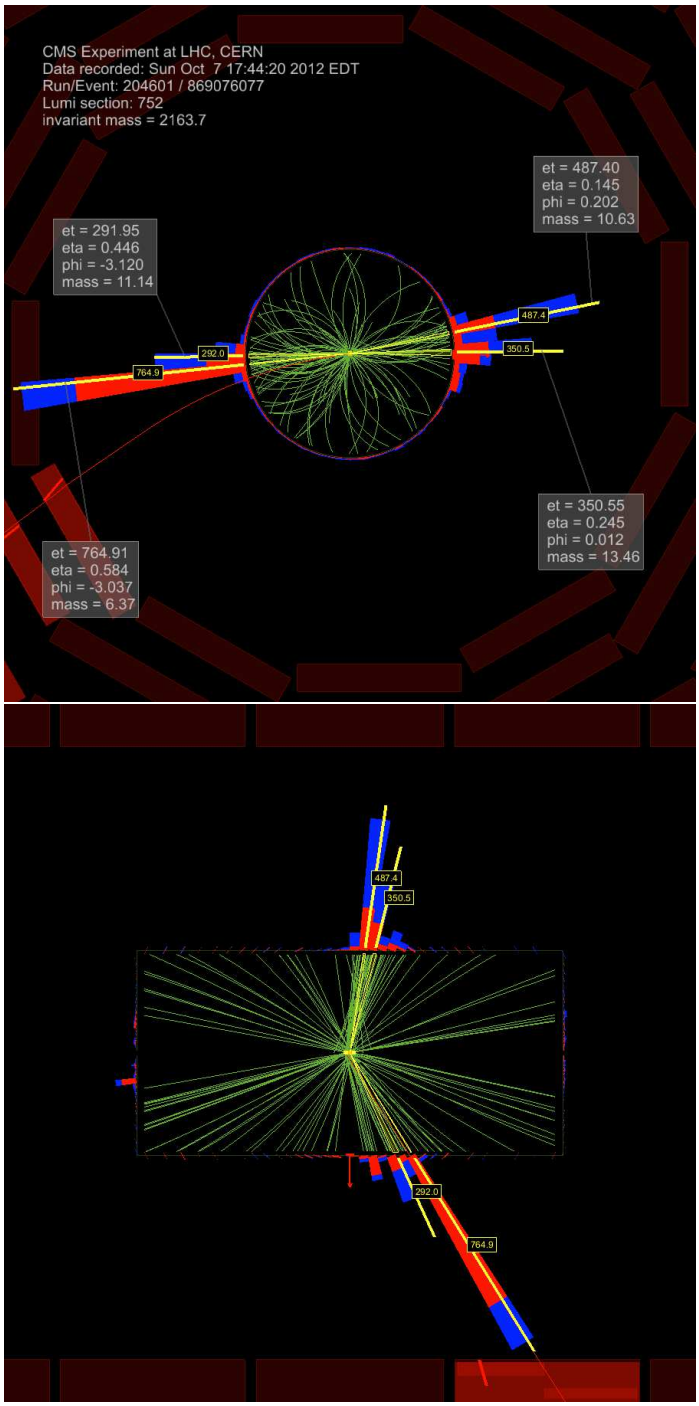

Figure A.4: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

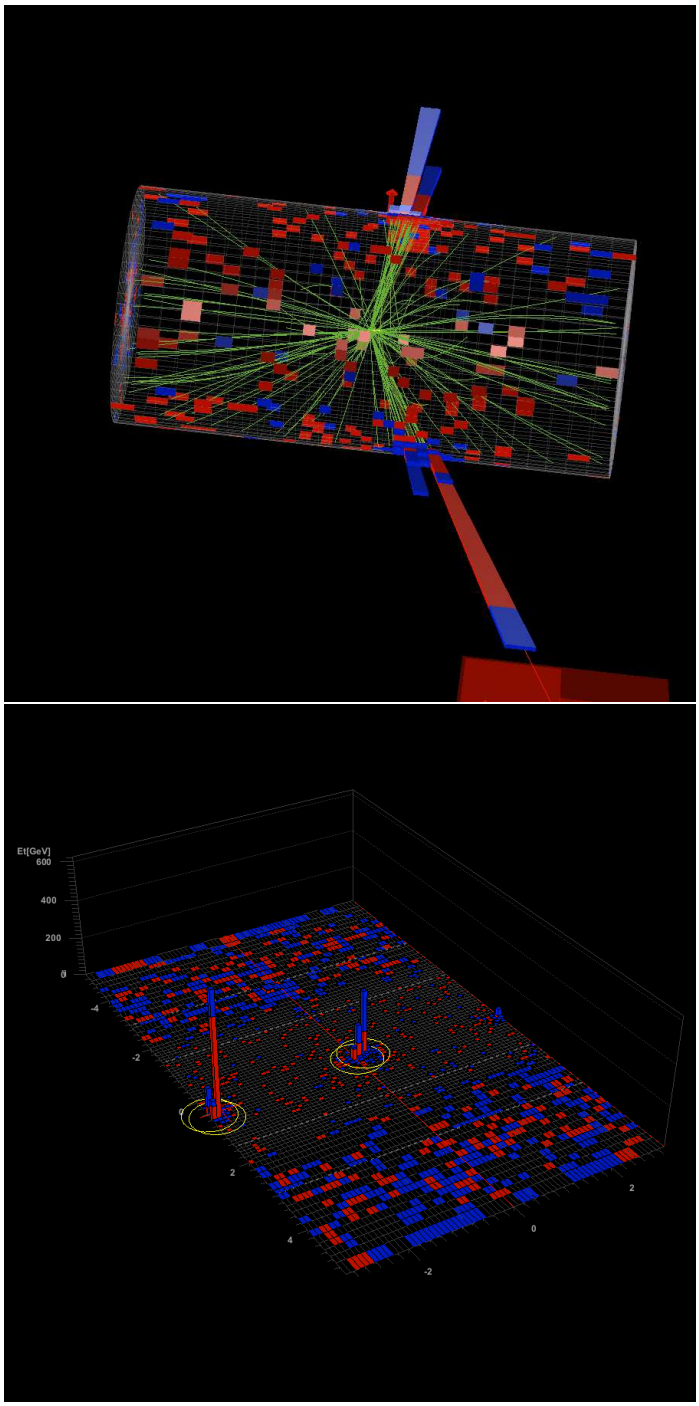

Figure A.5: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV . The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV . The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV .

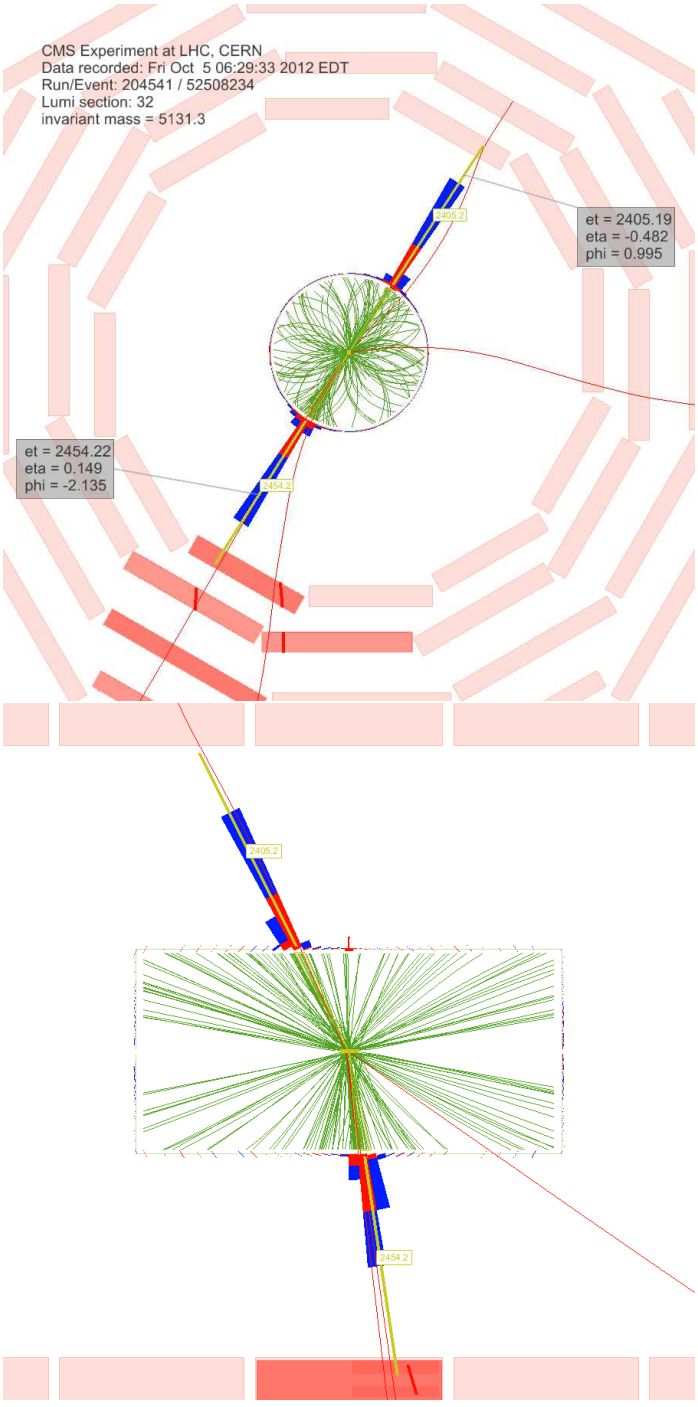

Figure A.6: Event display of event with the highest dijet invariant mass of 5.13 TeV . The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV .

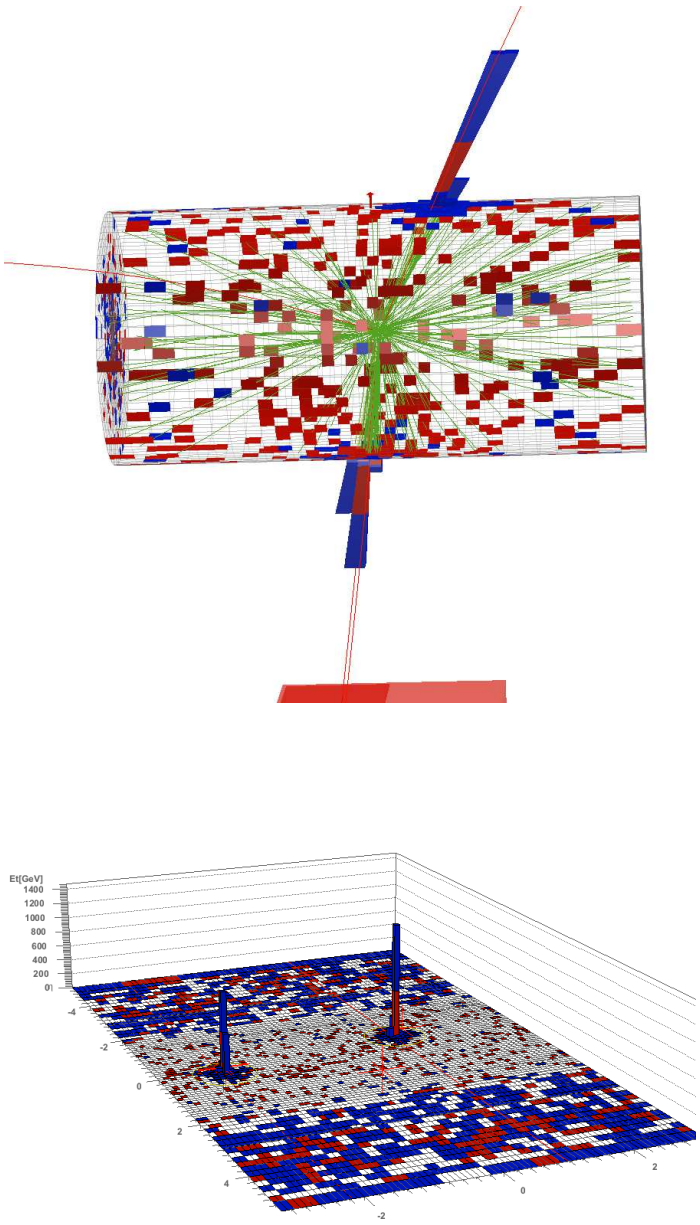

Figure A.7: Event display of event with the highest dijet invariant mass of 5.13 TeV . The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV .

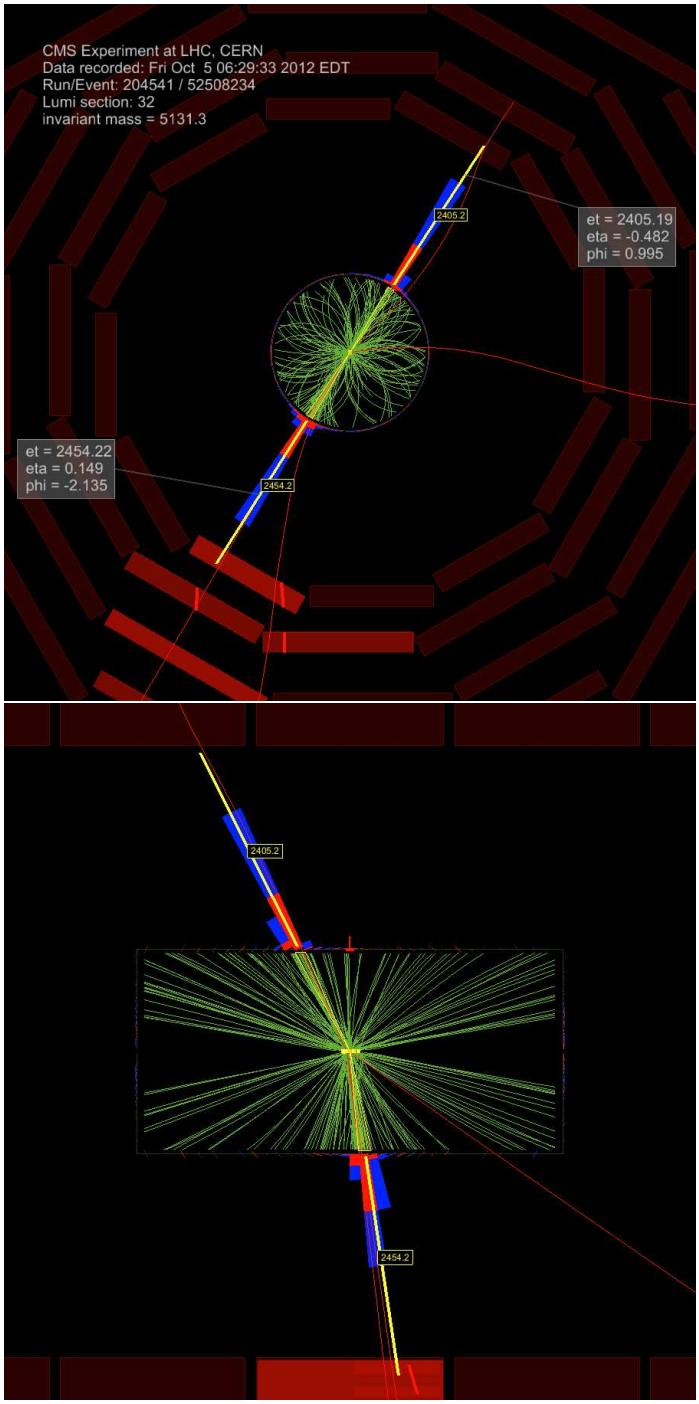

Figure A.8: Event display of event with the highest dijet invariant mass of 5.13 TeV . The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV .

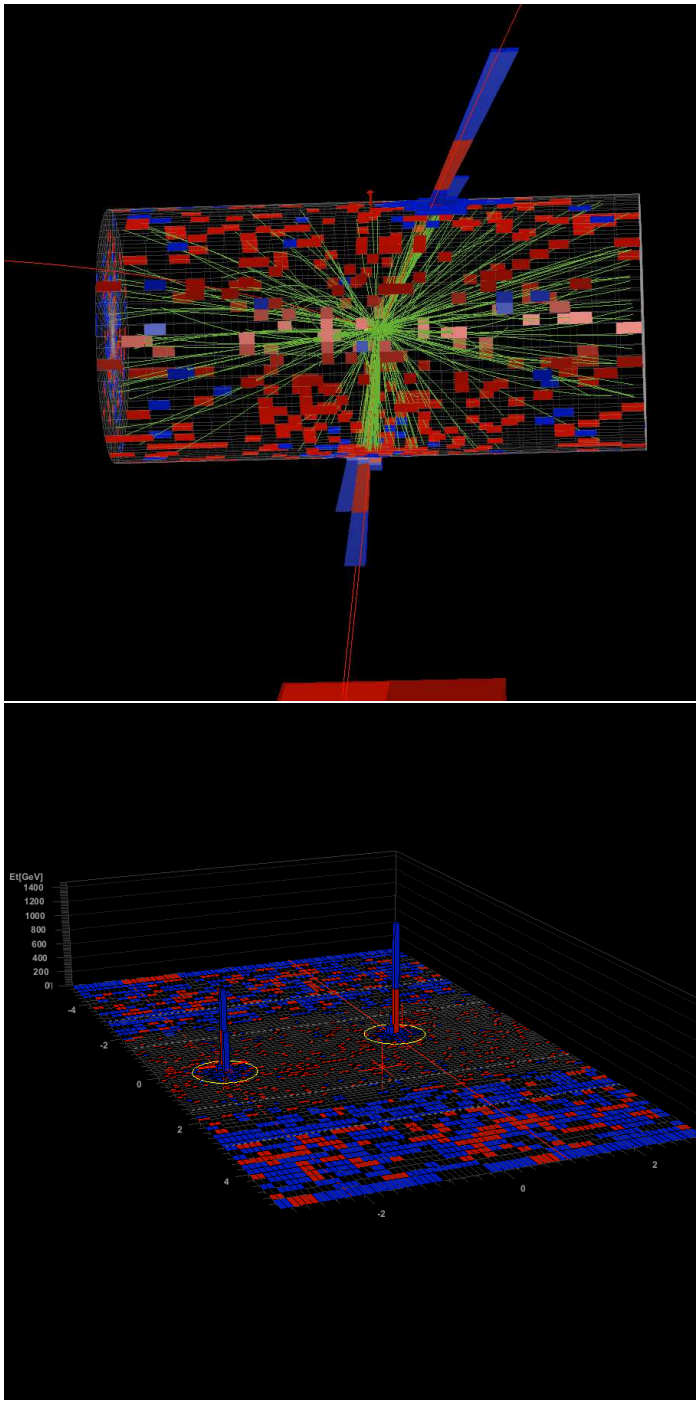

Figure A.9: Event display of event with the highest dijet invariant mass of 5.13 TeV . The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV .

A.3 Limit calculation cross check

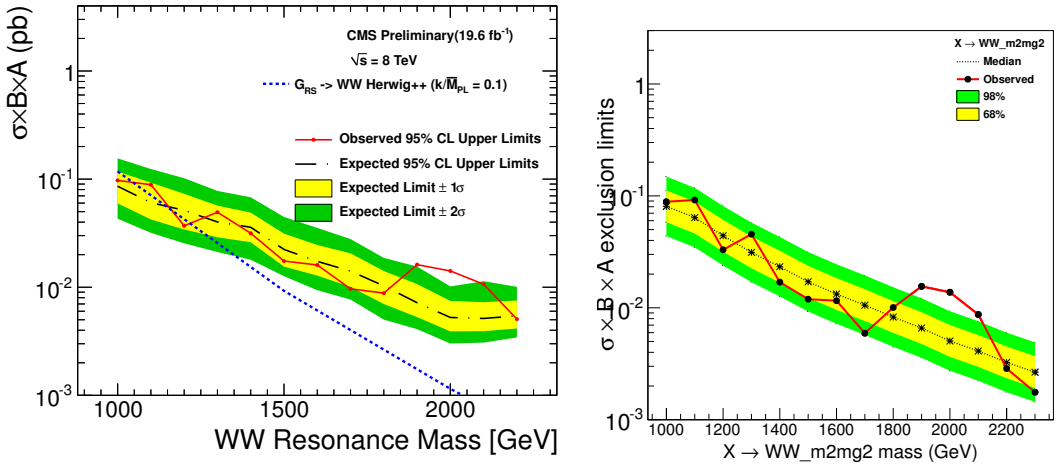

Figure A.10: Expected and observed limits on WW resonances in the 2-tag category. Left: Bayesian type limits as explained in section. Right: Asymptotic CLs type limits.

Appendix B

Appendices for Chapter 4

B.1 Model parameters and cross sections

We use scenario B of HVT model, in which we set $c_q = c_l = c_3 = 1.024$, $c_H = -0.976$, $g_V = 3$, $c_{VVV} = 0.928$, $c_{VW} = 1$, $c_{VVHH} = -0.024$. Other parameters are listed in Table B.1. In this table, we show the width of the generated W' and Z', also their production cross sections in different resonance masses.

Table B.1: Table of model parameters. CX+ is the cross section for W'+. CX- is the cross section for W'-. CX0 is the cross section for Z'. M is the resonance mass, in unit TeV. Wd_W′ and Wd_Z′ are the width of the W' and Z' signals, in unit of GeV.

M	g	Wd_Z'	$\mathcal{B}$(ZH)	Wd_W'	$\mathcal{B}$(WH)	CX+(pb)	CX0(pb)	CX-(pb)
1.3	0.645	42.492	0.496	42.467	0.490	8.41E-02	5.19E-02	2.56E-02
1.4	0.645	45.200	0.492	45.181	0.487	5.86E-02	3.55E-02	1.72E-02
1.5	0.646	47.961	0.489	47.946	0.485	4.10E-02	2.44E-02	1.16E-02
1.6	0.646	50.762	0.487	50.749	0.483	2.88E-02	1.68E-02	7.89E-03
1.8	0.647	56.447	0.484	56.438	0.481	1.42E-02	8.16E-03	3.71E-03
1.9	0.647	59.320	0.483	59.313	0.480	1.00E-02	5.71E-03	2.56E-03
2.0	0.647	62.209	0.482	62.203	0.479	7.10E-03	4.01E-03	1.78E-03
2.3	0.647	70.946	0.479	70.942	0.477	2.50E-03	1.41E-03	6.02E-04
2.4	0.647	73.876	0.479	73.873	0.477	1.76E-03	9.99E-04	4.21E-04
2.5	0.647	76.813	0.478	76.810	0.477	1.24E-03	7.08E-04	2.95E-04
2.6	0.647	79.756	0.478	79.753	0.476	8.70E-04	5.01E-04	2.07E-04
2.9	0.648	88.613	0.477	88.611	0.476	2.96E-04	1.78E-04	7.05E-05
3.0	0.648	91.573	0.477	91.571	0.475	2.05E-04	1.26E-04	4.91E-05

B.2 tau42 scale factor extrapolation

We extrapolate our $H \rightarrow \mathrm{WW}^* \rightarrow 4\mathrm{q}$ tagging scale factor, as in Equation B.1, from hadronic W tagging scale factor, as in Equation B.2, plus an additional systematic uncertainty ϵ, as in Equation B.3.

$$SF_H = \frac{\tau_{42}^{Data}}{\tau_{42}^{MC}} \tag{B.1}$$

$$SF_W = \frac{\tau_{21}^{Data}}{\tau_{21}^{MC}} \tag{B.2}$$

$$SF_H = SF_W + \epsilon \tag{B.3}$$

Derived from equation B.3, we have equation B.4.

$$SF_H = SF_W \Longleftrightarrow \frac{\tau_{42}^{Data}}{\tau_{42}^{MC}} = \frac{\tau_{21}^{Data}}{\tau_{21}^{MC}} \Longleftrightarrow \frac{\tau_{42}^{Data}}{\tau_{21}^{Data}} = \frac{\tau_{42}^{MC}}{\tau_{21}^{MC}} \tag{B.4}$$

We validate equation B.4 by comparing the $\frac{\tau_{42}}{\tau_{21}}$ in data and PYTHIAand HERWIGQCD MC, as shown in Figure B.1, B.2 and B.3. In this plot MC shows resonably good agreement with data.

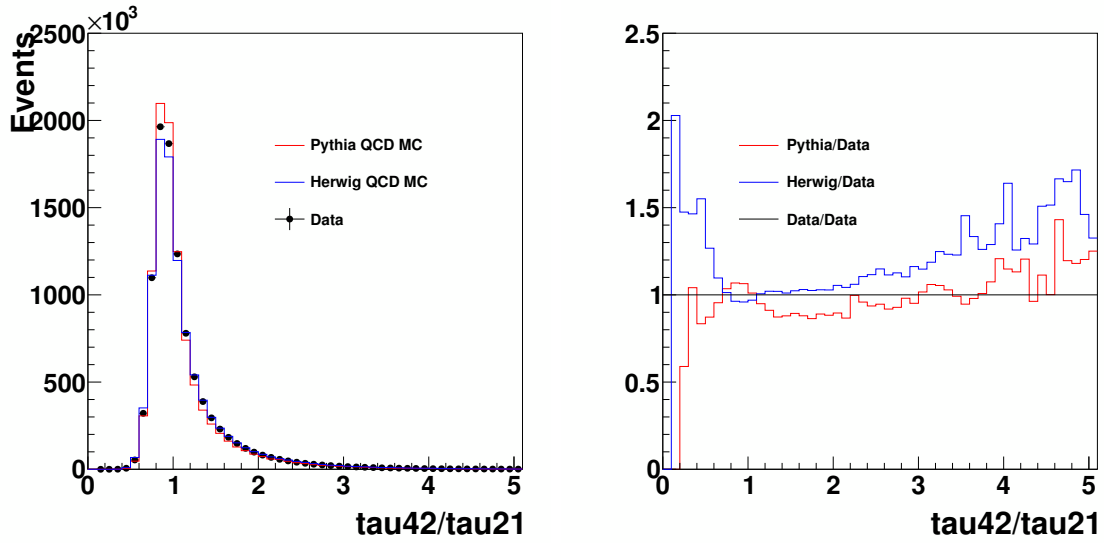

Figure B.1: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIAQCD MC (red), and HERWIGQCD MC (blue). Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

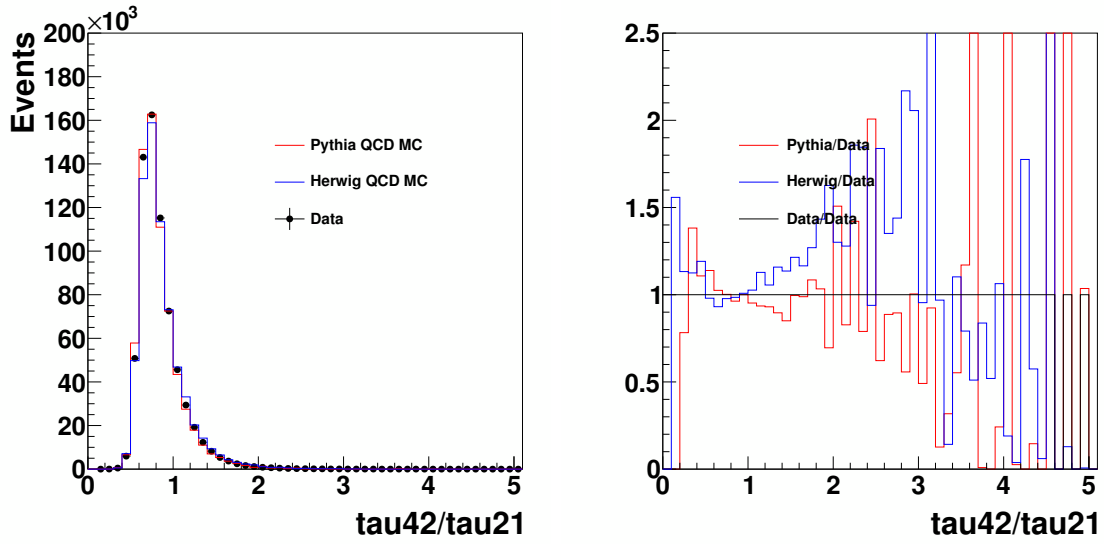

Figure B.2: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIAQCD MC (red), and HERWIGQCD MC (blue). Region of $\tau_{42} < 0.55$ is shown. Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

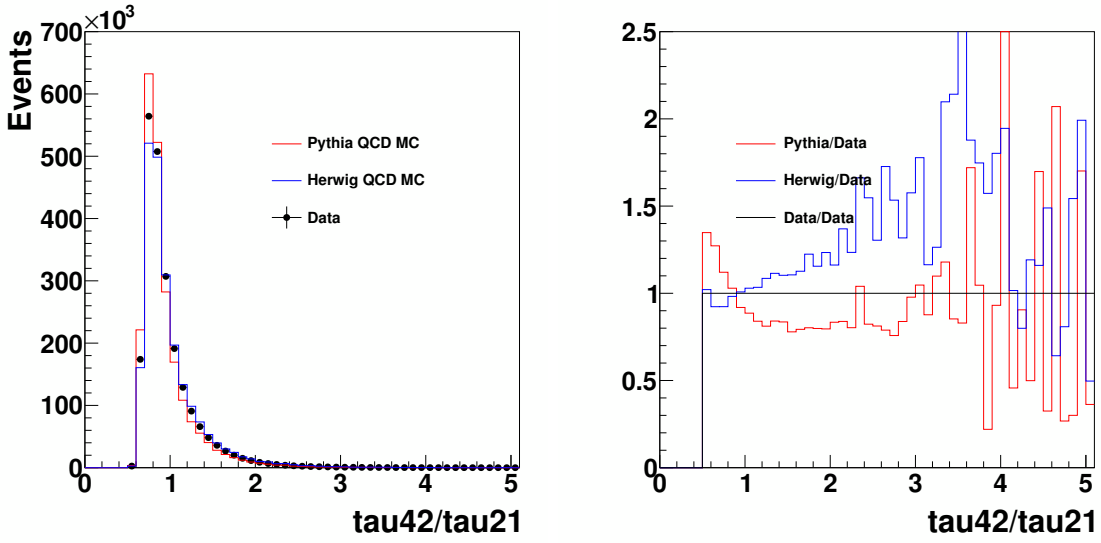

Figure B.3: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIAQCD MC (red), and HERWIGQCD MC (blue). Region of $0.55 < \tau_{42} < 0.65$ is shown. Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

Since Nsubjettiness τ_N is directly correlated with jet p_T, we further study the $\frac{\tau_{42}}{\tau_{21}}$ with respect to the jet p_T, which are shown in Figure B.4.

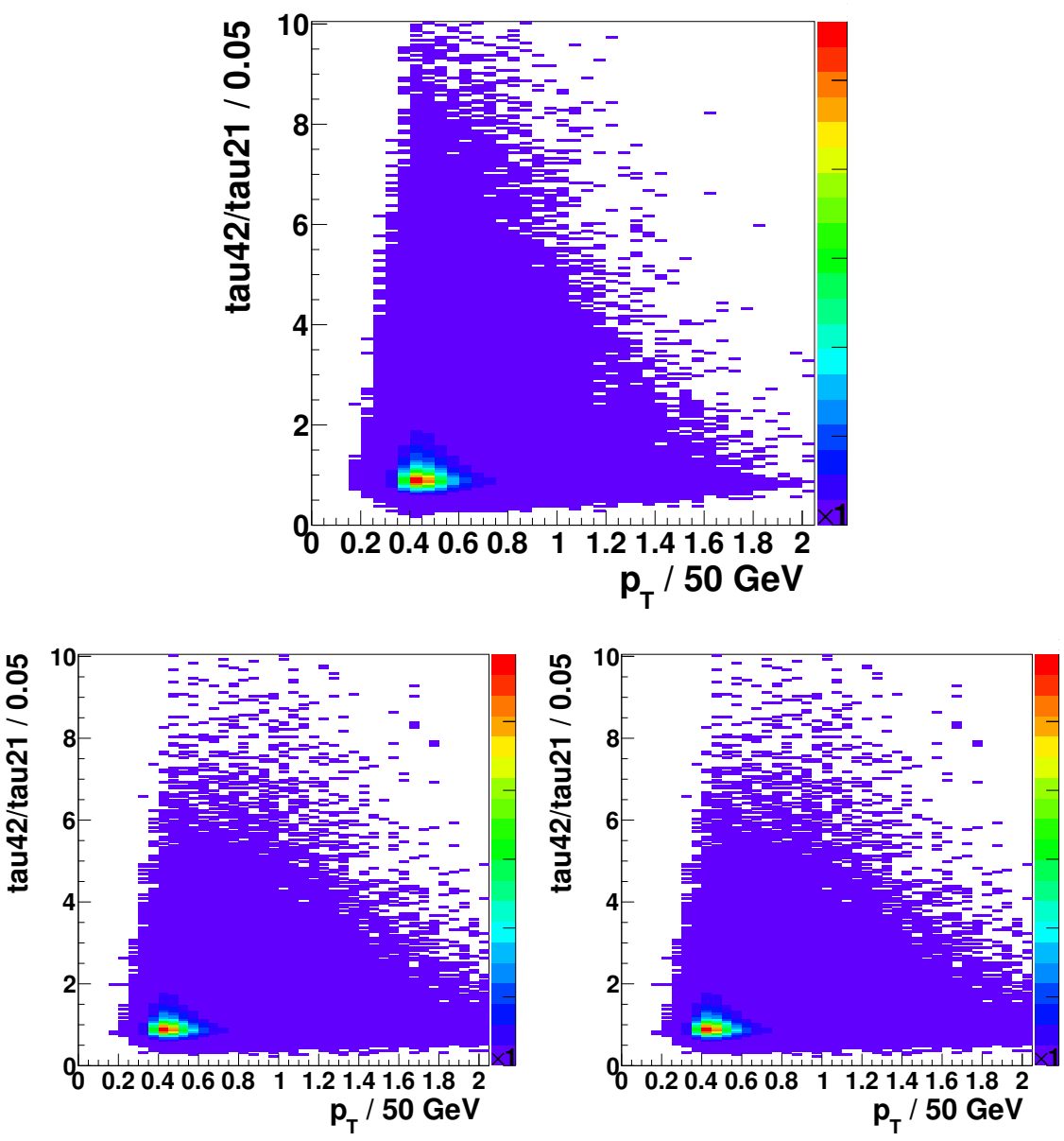

Figure B.4: 2D Plot of $\frac{\tau_{42}}{\tau_{21}}$ in Y axis, jet $p_T$in X axis, in data (top) compared to PYTHIAQCD MC (bottom left), and HERWIGQCD MC (bottom right).

The corresponding profile plot of Figure B.4 is shown in Figure B.5. From this plot, QCD MC agrees well with data. However, they still have a small discrepancy respect to data, and also between they seleves, especially at high p_T. To compensate this difference, we add an additional uncertainty ϵ, as mentioned in the beginning of this section. ϵ represents the shower and hadronization difference of MC tools, i.e., PYTHIA, HERWIG. We estimate ϵ by taking the largest difference in Higgs-tagging efficiency across various signal resonance masses, which results in 7%.

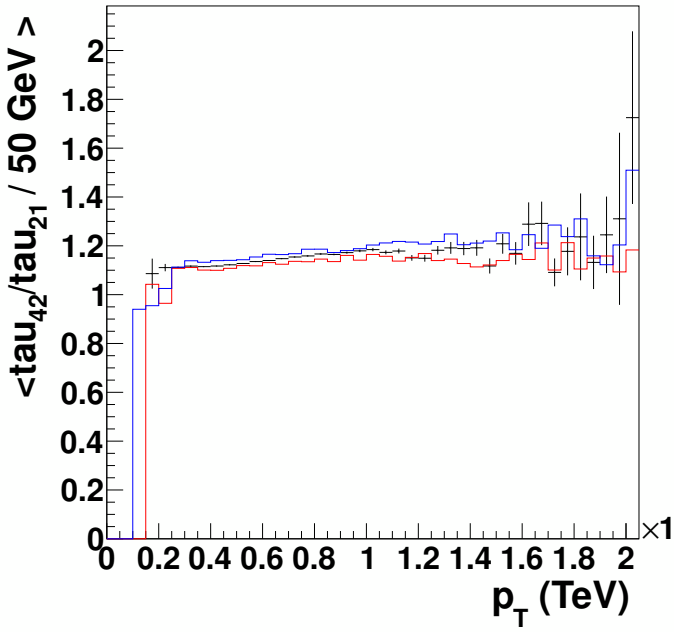

Figure B.5: Profile plot, mean of $\frac{\tau_{42}}{\tau_{21}}$ in Y axis, jet $p_T$in X axis, in data (black) compared to PYTHIAQCD MC (red), and HERWIGQCD MC (blue).

So the extrapolated scale factor for H-tagging is $0.86 \pm 7.6\% \pm 7.0\%$, for $\tau_{42} \leq 0.55$, while $1.39 \pm 54\% \pm 7.0\%$ for $0.55 < \tau_{42} < 0.65$.

B.3 CSVL Vs CSVM fat jet b tagging

In H(bb)Z analysis, We compare the csvl fat jet b tagging vs csvm on the limits. Fig.B.6 is showing the dijet mass spectrum with CSVL fat jet b tagging vs CSVM. And Fig.B.7 is showing the limits comparison between CSVL fat jet b tagging Vs CSVM. On the table, the limits from CSVL and CSVM are very close, so we show the limits in Table.B.2.

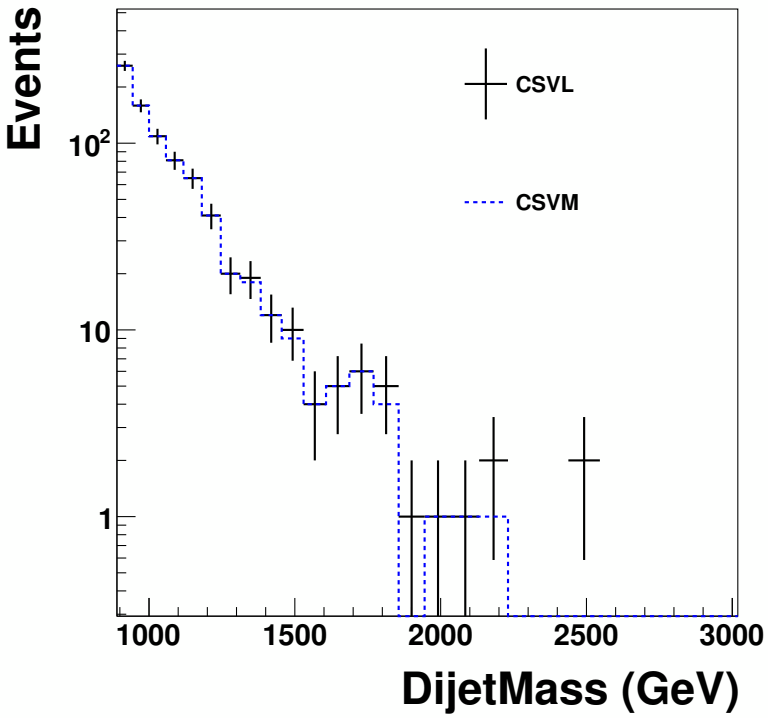

Figure B.6: DijetMass distribution for using CSVL Vs. CSVM.

Resonance(TeV)	CSVM(fb)	CSVL(fb)
1	29.9	29.4
1.5	6.86	6.98
2	2.98	2.67

Table B.2: Limits on different resonance mass for CSVL VS CSVM.

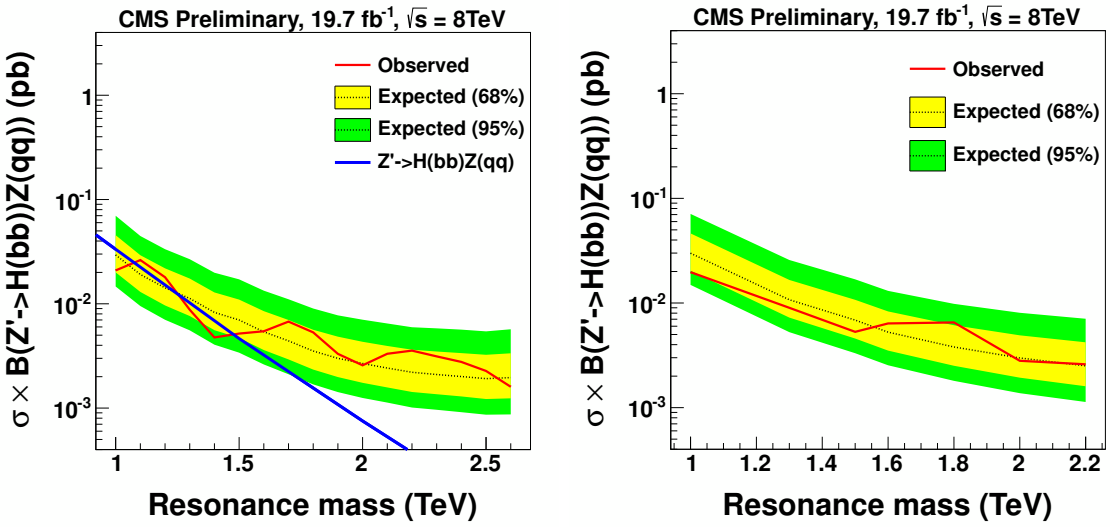

Figure B.7: Comparison for limits using CSVL fat jet b tagging(left), and CSVM fat jet b tagging(right).

B.4 tauNM distribution

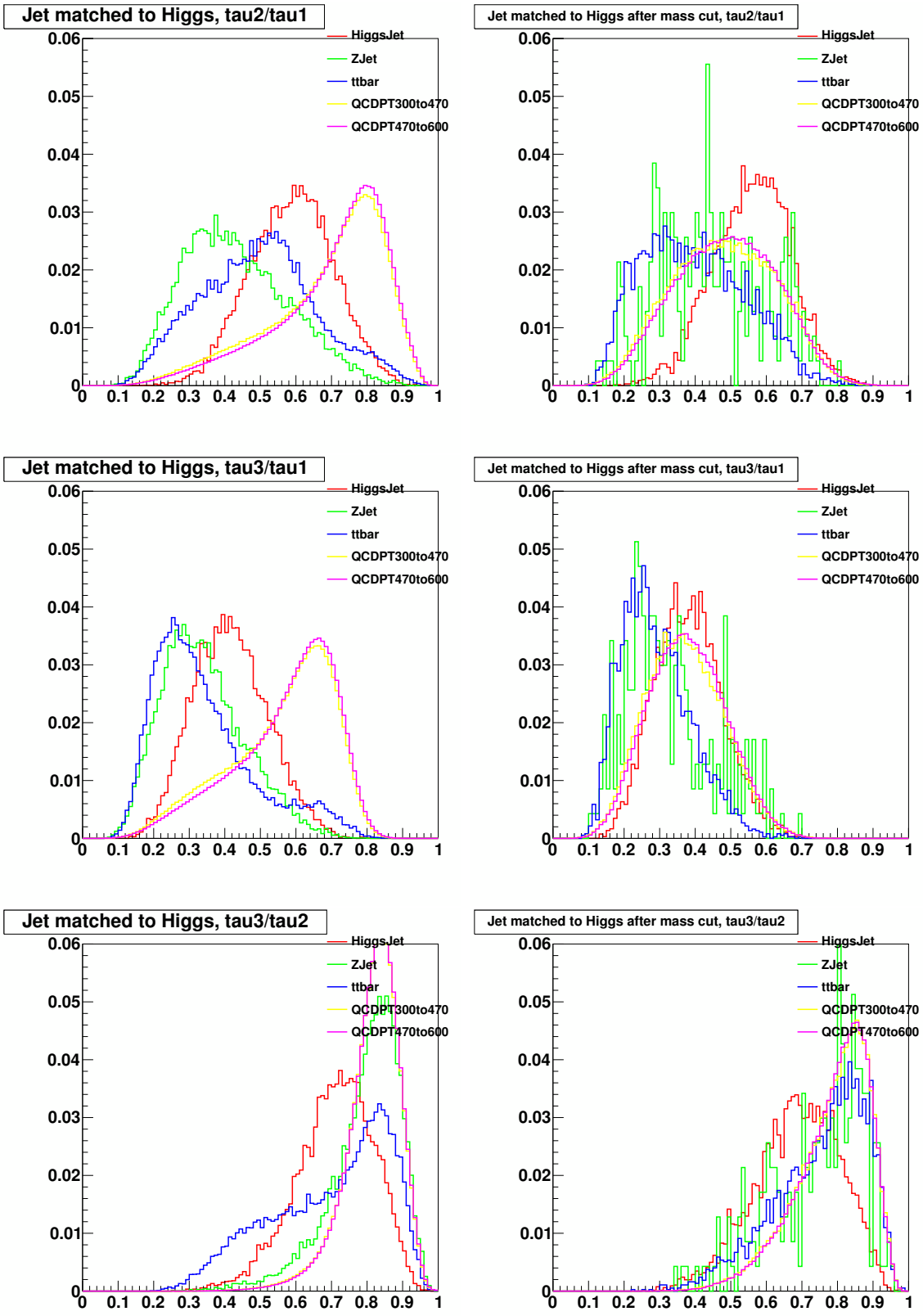

Figure B.8: list of tauNM plots between Higgs genJet and Z genJet, hadronic top and QCD. Signal used is 2 TeV Z'.

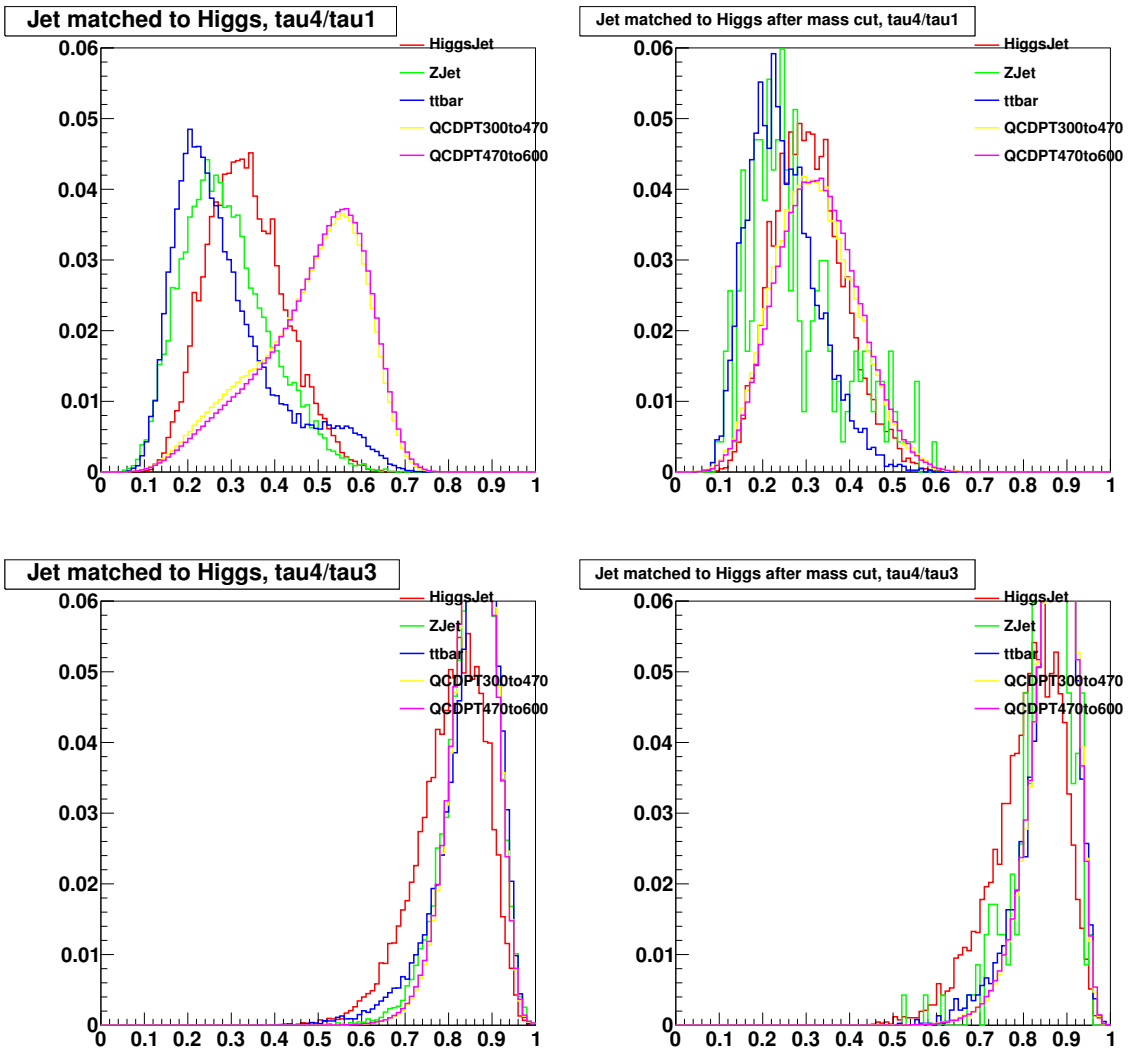

Figure B.9: list of tauNM plots between Higgs genJet and Z genJet, hadronic top and QCD. Signal used is 2 TeV Z'.

B.5 Cross-talk in data

The effect of the $H \to b\bar{b}$ tagger veto on the $H \to WW^* \to 4q$ tagged dijet mass distribution background (data) is shown in Figures B.10.

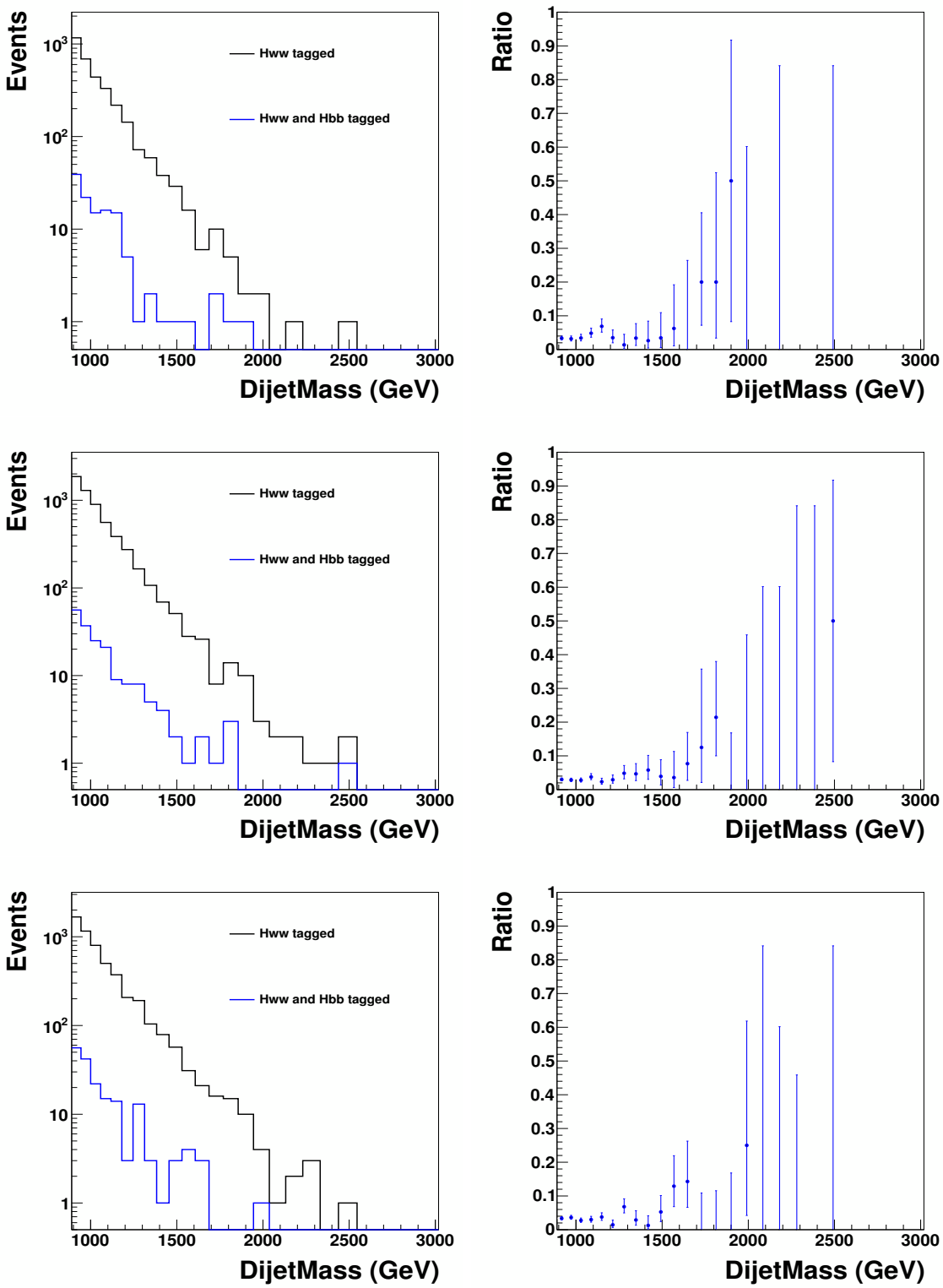

Figure B.10: Left column: dijet mass distribution in data, for events passing the $H \to WW^* \to 4q$ tagger (black), and for a subset of these events passing also the $H \to b\bar{b}$ tagger (blue). Right column: the fraction of $H \to WW^* \to 4q$ tagged events also tagged by $H \to b\bar{b}$. Top row: the high purity $H \to WW^* \to 4q$ tagger and high purity V-tagger. Middle row : the low purity $H \to WW^* \to 4q$ tagger, high purity V

Bibliography

[1] S.L.Glashow, *Nuclear Physics*, vol. 22, 1961.

[2] A.Salam, *Physics Review*, vol. 127, 1962.

[3] S. Weinberg, *Physics Review Letter*, vol. 19, 1967.

[4] C.Quigg, *Gauge Field Theories*. McGraw-Hill Book Company, 1983.

[5] Wikipedia, "Standard model," http://en.wikipedia.org/wiki/Standard_Model.

[6] David Griffiths, *Introduction to Elementary Particles*. Wiley-VCH Verlg GmbH & Co.KGaA, 2008.

[7] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, ser. Advanced book program. Boulder (Colo.): Westview Press Reading (Mass.), 1995, autre tirage : 1997. [Online]. Available: http://opac.inria.fr/record=b1131978

[8] M. Belloni, "qcd confinement illustration," http://www.phy.davidson.edu/FacHome/mjb/images/color_field.gif.

[9] NASA, "The accelerated expansion of universe," http://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy/.

[10] Quantum Diaries, "The hierarchy problem: why the higgs has a snowballs chance in hell," http://www.quantumdiaries.org/2012/07/01/the-hierarchy-problem-why-the-higgs-has-a-snowballs-chance-in-hell/.

[11] U. Baur, I. Hinchliffe, and D. Zeppenfeld, "Excited quark production at hadron colliders," *Int. J. Mod. Phys. A*, vol. 2, p. 1285, 1987.

[12] U. Baur, M. Spira, and P. M. Zerwas, "Excited-quark and -lepton production at hadron colliders," *Phys. Rev. D*, vol. 42, p. 815, 1990.

[13] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.*, vol. 83, p. 3370, 1999.

[14] L. Randall and R. Sundrum, "An alternative to compactification," *Phys. Rev. Lett.*, vol. 83, p. 4690, 1999.

[15] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, "Warped gravitons at the LHC and beyond," *Phys. Rev. D*, vol. 76, p. 036006, 2007.

[16] O. Antipin, D. Atwood, and A. Soni, "Search for RS gravitons via $w_l w_l$ decays," *Phys. Lett. B*, vol. 666, p. 155, 2008.

[17] O. Antipin and A. Soni, "Towards establishing the spin of warped gravitons," *JHEP*, vol. 10, p. 018, 2008.

[18] G. Altarelli, B. Mele, and M. Ruiz-Altaba, "Searching for new heavy vector bosons in $p\bar{p}$ colliders," *Z. Phys. C*, vol. 45, p. 109, 1989, erratum-ibid. C **47** (1990) 676.

[19] M. Gouzevitch, A. Oliveira, J. Rojo, R. Rosenfeld, G. Salam *et al.*, "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production," 2013.

[20] LHC Collaboration, "The overview of lhc," http://www.atlas.ch/photos/atlas_photos/selected-photos/detector-site/surface/.

[21] Wikipedia, "Pseudorapidity," http://en.wikipedia.org/wiki/Pseudorapidity.

[22] CMS Collaboration, *Physics TDR Vol 1. Detector Performance and Software*, 2006.

[23] CMS, "Scheme of a drift tube," http://www.pd.infn.it/~conti/cms/mycosmictest.html.

[24] CMS Collaboration, "Scheme of the carbon strip chamber," http://cms.web.cern.ch/news/cathode-strip-chambers.

[25] CMS Collaboration, "Study of jet substructure in pp collisions at 7 TeV in CMS," CMS Physics Analysis Summary CMS-PAS-JME-10-013, 2010. [Online]. Available: http://cdsweb.cern.ch/record/1333700

[26] CMS Collaboration, "Identifying hadronically decaying vector bosons merged into a single jet," no. CMS-PAS-JME-13-006, 2013. [Online]. Available: http://cdsweb.cern.ch/record/1598864

[27] CMS Collaboration, "Search for heavy resonances in the w/z-tagged dijet mass spectrum in pp collisions at 7 tev," *Phys. Lett. B*, vol. 723, p. 280, 2013.

[28] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness," *JHEP*, vol. 1103, p. 015, 2011.

[29] CMS Collaboration, "Search for narrow resonances using the dijet mass spectrum in pp collisions at $\sqrt{s} = 8$ TeV," 2012, submitted to Phys. Rev. Lett.

[30] ATLAS Collaboration, "Atlas search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 1301, p. 029, 2013.

[31] R. M. Harris and K. Kousouris, "Searches for Dijet Resonances at Hadron Colliders," *Int.J.Mod.Phys.*, vol. A26, pp. 5005–5055, 2011.

[32] G. Aad *et al.*, "Search for new phenomena in photon+jet events collected in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Lett. B*, vol. 728, p. 562, 2014.

[33] F. Abe *et al.*, "Search for excited quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.*, vol. 72, p. 3004, 1994.

[34] V. M. Abazov *et al.*, "Search for a heavy resonance decaying into a Z+ jet final state in p$\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using the D0 detector," *Phys. Rev. D*, vol. 74, p. 011104, 2006.

[35] CMS Collaboration, "Search for anomalous production of highly boosted Z bosons decaying to dimuons in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Lett. B*, vol. 722, p. 28, 2013.

[36] CMS Collaboration, "Search for a narrow spin-2 resonance decaying to a pair of Z vector bosons in the semileptonic final state," *Phys. Lett. B*, vol. 718, p. 1208, 2013.

[37] G. Aad *et al.*, "Search for resonant diboson production in the ww/wz $\rightarrow l\nu jj$ decay channels with the atlas detector at $\sqrt{s} = 7$ TeV," *Phys. Rev. D*, vol. 87, p. 112006, 2013.

[38] G. Aad *et al.*, "Search for new particles decaying to ZZ using final states with leptons and jets with the ATLAS detector in $\sqrt{s} = 7$ TeV proton-proton collisions," *Phys. Lett. B*, vol. 712, p. 331, 2012.

[39] T. Aaltonen *et al.*, "Search for high-mass resonances decaying into ZZ in p$\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Rev. D*, vol. 85, p. 012008, 2012.

[40] S. Chatrchyan *et al.*, "Search for new physics in final states with a lepton and

missing transverse energy in pp collisions at the lhc," *Phys. Rev. D*, vol. 87, p. 072005, 2013.

[41] G. Aad *et al.*, "Search for a heavy gauge boson decaying to a charged lepton and a neutrino in 1 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *Phys. Lett. B*, vol. 705, p. 28, 2011.

[42] S. Chatrchyan *et al.*, "Search for a W′ or ρ_{TC} decaying to WZ in pp collisions at $\sqrt{s} = 7$ tev," *Phys. Rev. Lett.*, vol. 109, p. 141801, 2012.

[43] G. Aad *et al.*, "Search for resonant WZ production in the WZ to l nu l' l' channel in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS detector," *Phys. Rev. D*, vol. 85, p. 112012, 2012.

[44] CMS Collaboration, "Jet performance in pp collisions at $\sqrt{s} = 7$ TeV," *CMS Physics Analysis Summary*, vol. CMS PAS JME-10-003, 2010. [Online]. Available: http://cdsweb.cern.ch/record/1279362

[45] CMS Collaboration, "Particle–flow event reconstruction in CMS and performance for jets, taus, and $\text{E}_{\text{T}}^{\text{miss}}$," CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009. [Online]. Available: http://cdsweb.cern.ch/record/1194487

[46] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector," *CMS Physics*

Analysis Summary, vol. CMS PAS PFT-10-001, 2010. [Online]. Available: http://cdsweb.cern.ch/record/1247373

[47] S. Chatrchyan *et al.*, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in *pp* Collisions at $\sqrt{s} = 7$ TeV," *JINST*, vol. 8, p. P09009, 2013.

[48] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross sections in deep-inelastic scattering," 1998.

[49] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms," *JHEP*, vol. 08, p. 001, 1997.

[50] M. Cacciari and G. P. Salam, "Dispelling the N**3 myth for the k(t) jet-finder," *Phys. Lett. B*, vol. 641, p. 57, 2006.

[51] M. Cacciari, G. P. Salam, and G. Soyez, "Fastjet user manual," *Eur. Phys. J. C*, vol. 72, p. 1896, 2012.

[52] M. Cacciari, G. P. Salam, and G. Soyez, "The catchment area of jets," *JHEP*, vol. 04, p. 005, 2008.

[53] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys. Lett. B*, vol. 659, p. 119, 2008.

[54] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *J. Instrum.*, vol. 6, p. P11002, 2011.

[55] CMS Collaboration, "Jet energy corrections and uncertainties detector performance plots for 2012," *CMS Detector Performance Plots*, vol. DP-2012/012, 2012.

[56] CMS Collaboration, "A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging," CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009. [Online]. Available: http://cdsweb.cern.ch/record/1194489

[57] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure," *Phys. Rev. D*, vol. 80, p. 051501, 2009.

[58] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches," *Phys. Rev. D*, vol. 81, p. 094023, 2010.

[59] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness," *JHEP*, vol. 1202, p. 093, 2012.

[60] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "N-Jettiness: An Inclusive Event Shape to Veto Jets," *Phys.Rev.Lett.*, vol. 105, p. 092002, 2010.

[61] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze *et al.*, "Spin determination of single-produced resonances at hadron colliders," *Phys.Rev.*, vol. D81, p. 075022, 2010.

[62] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze *et al.*, "On the

spin and parity of a single-produced resonance at the LHC," *Phys.Rev.*, vol. D86, p. 095031, 2012.

[63] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual," *JHEP*, vol. 05, p. 026, 2006.

[64] S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Plätzer, P. Richardson, M. H. Seymour, A. Sherstnev, and B. R. Webber, "Herwig++ 2.5 release note."

[65] S. Agostinelli *et al.*, "GEANT4—a simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, p. 250, 2003.

[66] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*, vol. 07, p. 012, 2002.

[67] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "MRST2001: partons and alpha_s from precise deep inelastic scattering and tevatron jet data," *Eur. Phys. J. C*, vol. 23, p. 73, 2002.

[68] R. Field, "Early LHC underlying event data – findings and surprises," in *Proceeedings of the Hadron Collider Physics Symposium 2010*, 2010.

[69] CMS Collaboration, "Search for Randall-Sundrum graviton excitations in the CMS experiment," 2002.

[70] S. Chatrchyan, "Search for resonances in the dijet mass spectrum from 7 TeV pp collisions at CMS," *Phys. Lett. B*, vol. 704, p. 123, 2011.

[71] CMS Collaboration, "Status of the 8 tev jet energy corrections and uncertainties based on 11/fb of data in cms," *CMS Detector Performance Plots*, vol. DP-2013/011, 2013.

[72] CMS Collaboration, "Cms luminosity based on pixel cluster counting - summer 2013 update," CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013. [Online]. Available: http://cdsweb.cern.ch/record/1598864

[73] A. D. Martin, W. J. Stirling, R. S. Throne, and G. Watt, "Update of parton distributoins at nnlo," *Phys. Lett. B*, vol. 652, p. 292, 2007.

[74] G. Cowan, K. Cranmer, E. Gross, O. Vitells, "Asymptotic formulae for likelihood based tests of new physics," *Eur. Phys. J.*, vol. C71, p. 1554, 2011.

[75] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys.*, vol. G28, pp. 2693–2704, 2002.

[76] T. Junk, "Confidence level computation for combining searches with small statistics," *Nucl. Instrum. Meth. A*, vol. 434, p. 435, 1999.

[77] "Procedure for the LHC Higgs boson search combination in Summer 2011," CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

[78] B. Bellazzini, C. Csaki, and J. Serra, "Composite higgses," 2014.

[79] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology," *JHEP*, vol. 1110, p. 081, 2011.

[80] D. Marzocca, M. Serone, and J. Shu, "General Composite Higgs Models," *JHEP*, vol. 1208, p. 013, 2012.

[81] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little higgs model," *Phys. Rev. D*, vol. 67, p. 095004, 2003.

[82] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy Vector Triplets: Bridging Theory and Data," 2014.

[83] V. Khachatryan *et al.*, "Search for new resonances decaying via WZ to leptons in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Phys. Lett. B*, vol. 740, p. 83, 2015.

[84] G. Aad *et al.*, "Search for WZ resonances in the fully leptonic channel using pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Lett. B*, vol. 737, p. 223, 2014.

[85] G. Aad *et al.*, "Search for resonant diboson production in the $\ell\ell q\bar{q}$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 75, p. 69, 2015.

[86] V. Khachatryan *et al.*, "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 08, p. 173, 2014.

[87] V. Khachatryan *et al.*, "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 08, p. 174, 2014.

[88] V. Khachatryan *et al.*, "Search for narrow high-mass resonances in proton-proton collisions at sqrt(s) = 8 TeV decaying to Z and Higgs bosons," 2015, submitted to Phys. Lett. B.

[89] G. Aad *et al.*, "Search for a new resonance decaying to a W or Z boson and a Higgs boson in the $\ell\ell/\ell\nu/\nu\nu + b\bar{b}$ final states with the ATLAS Detector," 2015.

[90] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett.*, vol. B716, pp. 1–29, 2012.

[91] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012.

[92] V. Khachatryan *et al.*, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV," 2014.

[93] V. Khachatryan *et al.*, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV," 2014.

[94] The ATLAS collaboration, "Study of the spin and parity of the Higgs boson in HVV decays with the ATLAS detector," 2015.

[95] CMS Collaboration, "Performance of b tagging at $\sqrt{s} = 8$ tev in multijet, $t\bar{t}$ and boosted topology events," *CMS PAS BTV-13-001*, 2013.

[96] C. Collaboration, "Search for narrow resonances using the dijet mass spectrum with 19.6fb-1 of pp collisions at sqrts=8 tev," *CMS Physics Analysis Summary*, vol. CMS-PAS-EXO-12-059, 2013. [Online]. Available: http://cds.cern.ch/record/1519066?ln=en

[97] C. Collaboration, "Search for heavy resonances decaying into two higgs bosons in hadronic final states," *CMS Physics Analysis Summary*, vol. CMS-PAS-EXO-12-053, 2013.

[98] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, "MadGraph 5: Going beyond," *JHEP*, 2011.

[99] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.

[100] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Resources for heavy vector triplets," https://rtorre.web.cern.ch/rtorre/Riccardotorre/heptools.html.

[101] M. Carena, C. Grojean, M. Kado, and V. Sharma, "Status of higgs boson physics," *PDG*, 2013. [Online]. Available: http://pdg.lbl.gov/2013/reviews/rpp2013-rev-higgs-boson.pdf

[102] C. Collaboration, "Search for a bsm resonance decaying into w and higgs bosons in lvbb final state," *CMS Physics Analysis Summary*, vol. CMS-PAS-EXO-14-010, 2014.

[103] M. B. D.E.Kaplan, K.Rehermann, "Top tagging: A method for identifying boosted hadronically decaying top quarks," *Phys. Rev. Lett*, vol. 101, p. 142001, 2008.

[104] R. G. Lomax and D. L. Hahs-Vaughn, *Statistical Concepts: A Second Course*. p.10, Routledge Academic, 2007.

[105] V. Khachatryan *et al.*, "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 08, p. 173, 2014.

[106] G. Punzi, "Sensitivity of Searches for New Signals and Its Optimization," in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, L. Lyons, R. Mount, and R. Reitmeyer, Eds., 2003, p. 79.

[107] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky *et al.*, "New parton distributions for collider physics," *Phys. Rev. D*, vol. 82, p. 074024, 2010.

[108] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur. Phys. J. C*, vol. 63, p. 189, 2009.

[109] R. D. Ball *et al.*, "Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties," *JHEP*, vol. 05, p. 075, 2010.

[110] S. Alekhin, S. Alioli, R. D. Ball, V. Bertone, J. Blumlein *et al.*, "The PDF4LHC Working Group Interim Report," 2011.

Vita

Education

- Ph.D. Physics, Johns Hopkins University, Expected 2015.
- M.A. Physics, Johns Hopkins University, 2013.
- B.S. Physics,University of Science and Technology of China, 2010.

Experience

- Research
  - **CMS Collaboration**, 2010–Present
- Teaching
  - Johns Hopkins University, 2010–2015
- Outreach

- USA Science and Engineering Festival, 2012,2014
- JHU Physics Fair, 2011,2012,2013,2014

Publications and Physics Analysis Summaries (in order of contribution)

- *Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at* $\sqrt{s}$*=8 TeV*, CMS Collaboration, **CMS Physics Analysis Summary**, CMS PAS EXO-14-009 2014 (submitted to JHEP)
- *Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at sqrts = 8 TeV*, CMS Collaboration, **JHEP**, 10.1007/JHEP08(2014)173 2014
- *Identification techniques for highly boosted W bosons that decay into hadrons*, CMS Collaboration, **JHEP**, 10.1007/JHEP12(2014)017 2014